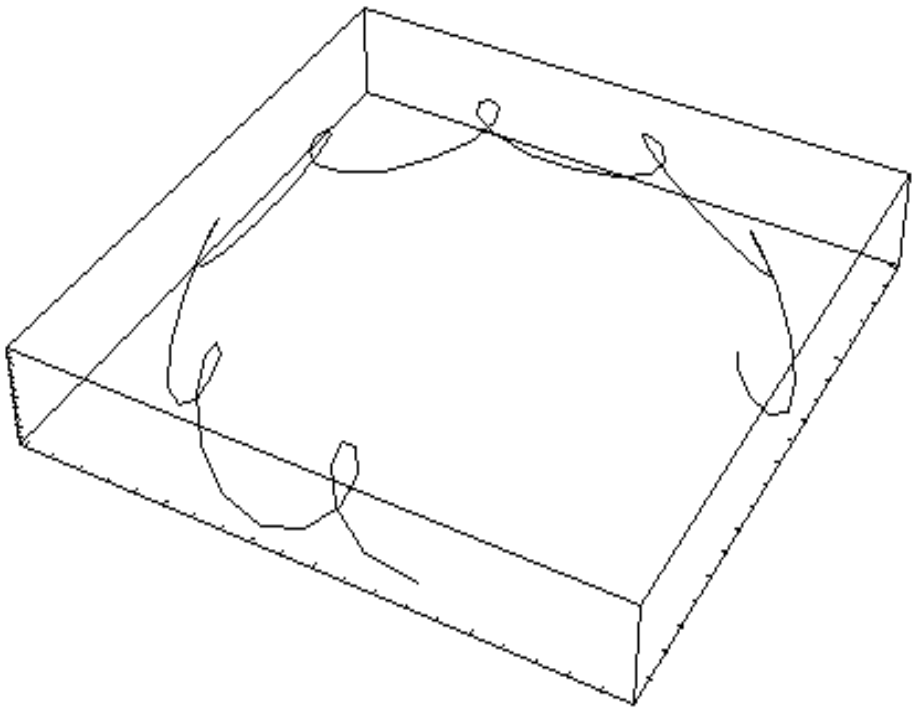
July 14, 2004

INTRODUCTION TO LAGRANGIAN AND HAMILTONIAN MECHANICS



Alain J. Brizard

Department of Chemistry and Physics
Saint Michael's College, Colchester, VT 05439

Contents

Chapter 1

Introduction to the Calculus of Variations

Minimum principles have been invoked throughout the history of Physics to explain the behavior of light and particles. In one of its earliest form, Heron of Alexandria (ca. 75 AD) stated that *light travels in a straight line* and that light follows a path of shortest distance when it is reflected by a mirror. In 1657, Pierre de Fermat (1601-1665) stated the Principle of *Least Time*, whereby light travels between two points along a path that minimizes the travel time, to explain Snell's Law (Willebrord Snell, 1591-1626) associated with light refraction in a stratified medium.

The mathematical foundation of the Principle of Least Time was later developed by Joseph-Louis Lagrange (1736-1813) and Leonhard Euler (1707-1783), who developed the mathematical method known as the *Calculus of Variations* for finding curves that minimize (or maximize) certain integrals. For example, the curve that maximizes the area enclosed by a contour of fixed length is the circle (e.g., a circle encloses an area $4/\pi$ times larger than the area enclosed by a square of equal perimeter length). The purpose of the present Chapter is to introduce the Calculus of Variations by means of applications of Fermat's Principle of Least Time.

1.1 Fermat's Principle of Least Time

According to Heron of Alexandria, light travels in a straight line when it propagates in a uniform medium. Using the index of refraction $n_0 \geq 1$ of the uniform medium, the speed of light in the medium is expressed as $v_0 = c/n_0 \leq c$, where c is the speed of light in vacuum. This straight path is not only a path of shortest distance but also a path of least time.

According to Fermat's Principle (Pierre de Fermat, 1601-1665), light propagates in a nonuniform medium by travelling along a path that *minimizes* the travel time between an

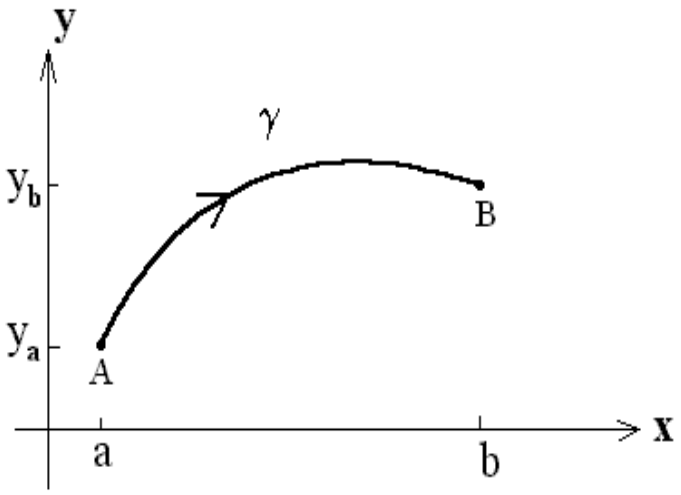


Figure 1.1: Light path in a nonuniform medium

initial point A (where a light ray is launched) and a final point B (where the light ray is received). Hence, the time taken by a light ray following a path γ from point A to point B (parametrized by σ) is

$$T_\gamma \;=\; \int_\gamma \, c^{-1} n(\mathbf{x}) \, \left| \frac{d\mathbf{x}}{d\sigma} \right| \, d\sigma \;=\; c^{-1} \, L_\gamma, \tag{1.1}$$

where L_γ represents the length of the *optical* path taken by light. In Sections 1 and 2 of the present Chapter, we consider ray propagation in two dimensions and return to general properties of ray propagation in Section 3.

For ray propagation in two dimensions (labeled x and y) in a medium with nonuniform refractive index $n(y)$, an arbitrary point $(x,\ y = y(x))$ along the light path γ is parametrized by the x-coordinate [i.e., $\sigma = x$ in Eq. (1.1)], which starts at point $A = (a, y_a)$ and ends at point $B = (b, y_b)$ (see Figure 1.1). Note that the path γ is now represented by the mapping $y : x \mapsto y(x)$. Along the path γ, the infinitesimal length element is $ds = \sqrt{1 + (y')^2}\; dx$ along the path $y(x)$ and the optical length

$$L[y] \;=\; \int_a^b \, n(y) \, \sqrt{1 + (y')^2} \, dx \tag{1.2}$$

is now a *functional* of y (i.e., changing y changes the value of the integral $L[y]$).

For the sake of convenience, we introduce the function

$$F(y, y'; \, x) \;=\; n(y) \, \sqrt{1 + (y')^2} \tag{1.3}$$

to denote the integrand of Eq. (1.2); here, we indicate an explicit dependence on x of $F(y, y'; x)$ for generality.

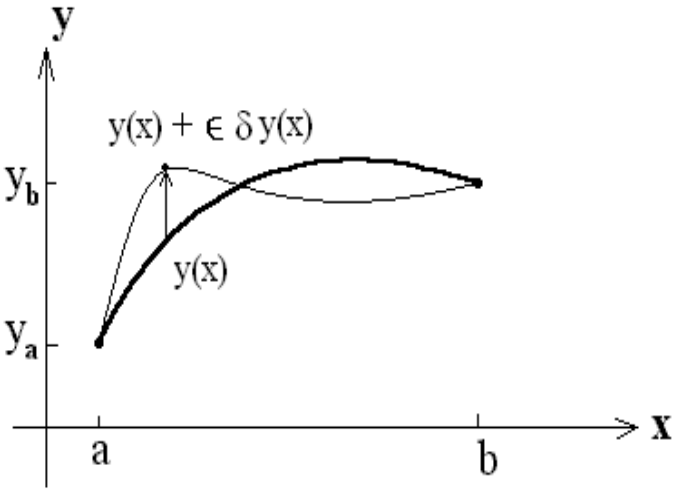


Figure 1.2: Virtual displacement

1.1.1 Euler's First Equation

We are interested in finding the curve $y(x)$ that minimizes the optical-path integral (1.2). The method of Calculus of Variations will transform the problem of minimizing an integral of the form $\int_a^b F(y, y'; x)\, dx$ into the solution of a differential equation expressed in terms of derivatives of the integrand $F(y, y'; x)$.

To determine the path of least time, we introduce the *functional* derivative $\delta L[y]$ defined as

$$\delta L[y] \;=\; \left(\frac{d}{d\epsilon} L[y + \epsilon\, \delta y]\right)_{\epsilon=0},$$

where $\delta y(x)$ is an arbitrary smooth variation of the path $y(x)$ subject to the boundary conditions $\delta y(a) = 0 = \delta y(b)$. Hence, the end points of the path are not affected by the variation (see Figure 1.2). Using the expression for $L[y]$ in terms of the function $F(y, y';\, x)$, we find

$$\delta L[y] \;=\; \int_a^b \left[\, \delta y(x)\, \frac{\partial F}{\partial y(x)} \;+\; \delta y'(x)\, \frac{\partial F}{\partial y'(x)} \,\right] dx,$$

where $\delta y' = (\delta y)'$, which when integrated by parts becomes

$$\delta L[y] \;=\; \int_a^b \delta y \left[\, \frac{\partial F}{\partial y} \;-\; \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \right] dx \;+\; \frac{1}{c}\left[\, \delta y_b \left(\frac{\partial F}{\partial y'}\right)_b \;-\; \delta y_a \left(\frac{\partial F}{\partial y'}\right)_a \right].$$

Here, since the variation $\delta y(x)$ vanishes at the integration boundaries ($\delta y_b = 0 = \delta y_a$), we obtain

$$\delta L[y] \;=\; \int_a^b \delta y \left[\, \frac{\partial F}{\partial y} \;-\; \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \right] dx. \tag{1.4}$$

The condition that the path γ takes the *least time*, corresponding to the variational principle $\delta L[y] = 0$, yields Euler's First equation

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \;=\; \frac{\partial F}{\partial y}. \tag{1.5}$$

This ordinary differential equation for $y(x)$ yields a solution that gives the desired path of least time.

We now apply the variational principle $\delta L[y] = 0$ for the case where F is given by Eq. (1.3), for which we find

$$\frac{\partial F}{\partial y'} \;=\; \frac{n(y)\,y'}{\sqrt{1+(y')^2}} \quad \text{and} \quad \frac{\partial F}{\partial y} \;=\; n'(y)\,\sqrt{1+(y')^2},$$

so that Euler's First Equation (1.5) becomes

$$n(y)\,y'' \;=\; n'(y)\,\left[1+(y')^2\right]. \tag{1.6}$$

Although the solution of this (nonlinear) second-order ordinary differential equation is difficult to obtain for general functions $n(y)$, we can nonetheless obtain a qualitative picture of its solution by noting that y'' has the same sign as $n'(y)$. Hence, when $n'(y) = 0$ (i.e., the medium is spatially uniform), the solution $y'' = 0$ yields the straight line $y(x;\phi_0) = \tan\phi_0\, x$, where ϕ_0 denotes the initial launch angle (as measured from the horizontal axis). The case where $n'(y) > 0$ (or $n'(y) < 0$), on the other hand, yields a light path which is concave upwards (or downwards) as will be shown below.

We should point out that Euler's First Equation (1.5) results from the extremum condition $\delta L[y] = 0$, which does not necessarily imply that the Euler path $y(x)$ actually minimizes the optical length $L[y]$. To show that the path $y(x)$ minimizes the optical length $L[y]$, we must evaluate the second functional derivative

$$\delta^2 L[y] \;=\; \left(\frac{d^2}{d\epsilon^2} L[y+\epsilon\,\delta y]\right)_{\epsilon=0}.$$

By following steps similar to the derivation of Eq. (1.4), we find

$$\delta^2 L[y] \;=\; \int_a^b \left\{\, \delta y^2 \left[\, \frac{\partial^2 F}{\partial y^2} \;-\; \frac{d}{dx}\left(\frac{\partial^2 F}{\partial y \partial y'}\right) \,\right] \;+\; (\delta y')^2 \;\frac{\partial^2 F}{\partial (y')^2} \,\right\}.$$

The necessary and sufficient condition for a minimum is $\delta^2 L > 0$ and, thus, the sufficient conditions for a minimal optical length are

$$\frac{\partial^2 F}{\partial y^2} \;-\; \frac{d}{dx}\left(\frac{\partial^2 F}{\partial y \partial y'}\right) \;>\; 0 \quad \text{and} \quad \frac{\partial^2 F}{\partial (y')^2} \;>\; 0,$$

for all smooth variations $\delta y(x)$. Using Eqs. (1.3) and (1.6), we find

$$\frac{\partial^2 F}{(\partial y')^2} \;=\; \frac{n}{[1+(y')^2]^{3/2}} \;>\; 0$$

and

$$\frac{\partial^2 F}{\partial y^2} \;=\; n''\,\sqrt{1+(y')^2} \quad \text{and} \quad \frac{\partial^2 F}{\partial y\,\partial y'} \;=\; \frac{n'\,y'}{\sqrt{1+(y')^2}}$$

so that

$$\frac{\partial^2 F}{\partial y^2} \;-\; \frac{d}{dx}\left(\frac{\partial^2 F}{\partial y \partial y'}\right) \;=\; \frac{n^2}{F}\,\frac{d^2 \ln n}{dy^2}.$$

Hence, the sufficient condition for a minimal optical length for light traveling in a nonuniform refractive medium is $d^2 \ln n/dy^2 > 0$.

1.1.2 Euler's Second Equation

Under certain conditions, we may obtain a partial solution to Euler's First Equation (1.6) for a light path $y(x)$ in a nonuniform medium. This partial solution is provided by Euler's Second equation, which is derived as follows.

First, we write the exact derivative dF/dx for $F(y, y'; x)$ as

$$\frac{dF}{dx} \;=\; \frac{\partial F}{\partial x} \;+\; y'\,\frac{\partial F}{\partial y} \;+\; y''\,\frac{\partial F}{\partial y'},$$

and substitute Eq. (1.5) to combine the last two terms so that we obtain Euler's Second equation

$$\frac{d}{dx}\left(F \;-\; y'\,\frac{\partial F}{\partial y'}\right) \;=\; \frac{\partial F}{\partial x}. \tag{1.7}$$

In the present case, the function $F(y, y'; x)$, given by Eq. (1.3), is explicitly independent of x (i.e., $\partial F/\partial x = 0$), and we find

$$F \;-\; y'\,\frac{\partial F}{\partial y'} \;=\; \frac{n(y)}{\sqrt{1+(y')^2}} \;=\; \text{constant},$$

and thus the partial solution of Eq. (1.6) is

$$n(y) \;=\; \alpha\,\sqrt{1+(y')^2}, \tag{1.8}$$

where α is a constant determined from the initial conditions of the light ray; note that, since the right side of Eq. (1.8) is always greater than α, we find that $n(y) > \alpha$. This is indeed a *partial* solution (in some sense), since we have reduced the derivative order from

second-order derivative in Eq. (1.6) to first-order derivative in Eq. (1.8): $y''(x) \rightarrow y'(x)$ on the solution $y(x)$.

Euler's Second Equation has, thus, produced an equation of the form $G(y, y'; x) = 0$, which can normally be integrated by *quadrature*. Here, Eq. (1.8) can be integrated by quadrature to give the integral solution

$$x(y) \;=\; \int_0^y \frac{\alpha \; d\eta}{\sqrt{[n(\eta)]^2 \;-\; \alpha^2}}, \tag{1.9}$$

subject to the condition $x(y = 0) = 0$. From the explicit dependence of the index of refraction $n(y)$, one may be able to perform the integration in Eq. (1.9) to obtain $x(y)$ and, thus, obtain an explicit solution $y(x)$ by inverting $x(y)$.

For example, let us consider the path associated with the index of refraction $n(y) = H/y$, where the height H is a constant and $0 < y < H\,\alpha^{-1}$ to ensure that, according to Eq. (1.8), $n(y) > \alpha$. The integral (1.9) can then be easily integrated to yield

$$x(y) \;=\; \int_0^y \frac{\alpha \; \eta \; d\eta}{\sqrt{H^2 \;-\; (\alpha\,\eta)^2}} \;=\; H\,\alpha^{-1} \left(1 \;-\; \sqrt{1 \;-\; \left(\frac{\alpha\, y}{H} \right)^2} \right).$$

Hence, the light path simply forms a semi-circle of radius $R = \alpha^{-1}\, H$ centered at $(x, y) = (R, 0)$:

$$(R - x)^2 + y^2 \;=\; R^2 \quad \rightarrow \quad y(x) \;=\; \sqrt{x\,(2R - x)}.$$

The light path is indeed concave downward since $n'(y) < 0$.

Returning to Eq. (1.8), we note that it states that as a light ray enters a region of increased (decreased) refractive index, the slope of its path also increases (decreases). In particular, by substituting Eq. (1.6) into Eq. (1.8), we find

$$\alpha^2 \; y'' \;=\; \frac{1}{2} \; \frac{dn^2(y)}{dy},$$

and, hence, the path of a light ray is concave upward (downward) where $n'(y)$ is positive (negative), as previously discussed.

1.1.3 Snell's Law

Let us now consider a light ray travelling in two dimensions from $(x, y) = (0, 0)$ at an angle ϕ_0 (measured from the x-axis) so that $y'(0) = \tan\phi_0$ is the slope at $x = 0$, assuming that $y(0) = 0$. The constant α is then simply determined from initial conditions as

$$\alpha \;=\; n_0 \; \cos\phi_0,$$

where $n_0 = n(0)$ is the refractive index at $y(0) = 0$. Next, let $y'(x) = \tan\phi(x)$ be the slope of the light ray at $(x,\, y(x))$, then $\sqrt{1+(y')^2} = \sec\phi$ and Eq. (1.8) becomes $n(y)\,\cos\phi = n_0\,\cos\phi_0$, which, when we substitute the complementary angle $\theta = \pi/2 - \phi$, finally yields the standard form of Snell's Law:

$$n[y(x)]\ \sin\theta(x)\ =\ n_0\ \sin\theta_0, \tag{1.10}$$

properly generalized to include a light path in a nonuniform refractive medium. Note that Snell's Law does not tell us anything about the actual light path $y(x)$; this solution must come from solving Eq. (1.9).

1.1.4 Application of Fermat's Principle

As an application of the Principle of Least Time, we consider the propagation of a light ray in a medium with refractive index $n(y) = n_0\ (1 - \beta\, y)$ exhibiting a constant gradient $n'(y) = -\,n_0\,\beta$.

Once again with $\alpha = n_0\,\cos\phi_0$ and $y'(0) = \tan\phi_0$, Eq. (1.8) becomes

$$1\ -\ \beta\, y\ =\ \cos\phi_0\,\sqrt{1+(y')^2}.$$

By separating dy and dx we obtain

$$dx\ =\ \frac{\cos\phi_0\ dy}{\sqrt{(1-\beta\, y)^2 - \cos^2\phi_0}}. \tag{1.11}$$

We now use the trigonometric substitution

$$1 - \beta\, y = \cos\phi_0\ \sec\theta, \tag{1.12}$$

with $\theta = \phi_0$ at $y = 0$, to find

$$dy\ =\ -\,\frac{\cos\phi_0}{\beta}\ \sec\theta\,\tan\theta\ d\theta$$

and

$$\sqrt{(1-\beta\, y)^2 - \cos^2\phi_0}\ =\ \cos\phi_0\ \tan\theta,$$

so that Eq. (1.11) becomes

$$dx\ =\ -\,\frac{\cos\phi_0}{\beta}\ \sec\theta\ d\theta. \tag{1.13}$$

The solution to this equation, with $x = 0$ when $\theta = \phi_0$, is

$$x\ =\ -\,\frac{\cos\phi_0}{\beta}\ \ln\left(\frac{\sec\theta + \tan\theta}{\sec\phi_0 + \tan\phi_0}\right). \tag{1.14}$$

If we can now solve for $\sec\theta$ as a function of x from Eq. (1.14), we can substitute this solution into Eq. (1.12) to obtain an expression for the light path $y(x)$. For this purpose, we define

$$\psi \;=\; \frac{\beta\, x}{\cos\phi_0} \;-\; \ln(\sec\phi_0 + \tan\phi_0),$$

so that Eq. (1.14) becomes

$$\sec\theta \;+\; \sqrt{\sec^2\theta \;-\; 1} \;=\; e^{-\psi},$$

which can be solved for $\sec\theta$ as

$$\sec\theta \;=\; \cosh\psi \;=\; \cosh\left[\frac{\beta\, x}{\cos\phi_0} \;-\; \ln(\sec\phi_0 + \tan\phi_0)\right].$$

Substituting this equation into Eq. (1.12), we find the light path

$$y(x;\beta) \;=\; \frac{1}{\beta} \;-\; \frac{\cos\phi_0}{\beta}\; \cosh\left[\frac{\beta\, x}{\cos\phi_0} \;-\; \ln(\sec\phi_0 + \tan\phi_0)\right]. \tag{1.15}$$

Note that, using the identities

$$\left.\begin{aligned} \cosh\left[\ln(\sec\phi_0 + \tan\phi_0)\right] &= \sec\phi_0 \\ \sinh\left[\ln(\sec\phi_0 + \tan\phi_0)\right] &= \tan\phi_0 \end{aligned}\right\}, \tag{1.16}$$

we can check that, in the uniform case ($\beta = 0$), we recover the expected result

$$\lim_{\beta\to 0}\; y(x;\beta) \;=\; (\tan\phi_0)\; x.$$

Next, we observe that $y(x;\beta)$ exhibits a single maximum located at $x = \overline{x}(\beta)$. Solving for $\overline{x}(\beta)$ from $y'(\overline{x};\beta) = 0$, we obtain

$$\tanh\left(\frac{\beta\,\overline{x}}{\cos\phi_0}\right) \;=\; \sin\phi_0,$$

or

$$\overline{x}(\beta) \;=\; \frac{\cos\phi_0}{\beta}\; \ln(\sec\phi_0 + \tan\phi_0), \tag{1.17}$$

and hence

$$y(x;\beta) \;=\; \frac{1}{\beta} \;-\; \frac{\cos\phi_0}{\beta}\; \cosh\left[\frac{\beta}{\cos\phi_0}\;(x - \overline{x})\right],$$

and $\overline{y}(\beta) = y(\overline{x};\beta) = (1 - \cos\phi_0)/\beta$. Figure 1.3 shows a graph of the normalized solution $y(x;\beta)/\overline{y}(\beta)$ as a function of the normalized coordinate $x/\overline{x}(\beta)$ for $\phi_0 = \frac{\pi}{3}$.

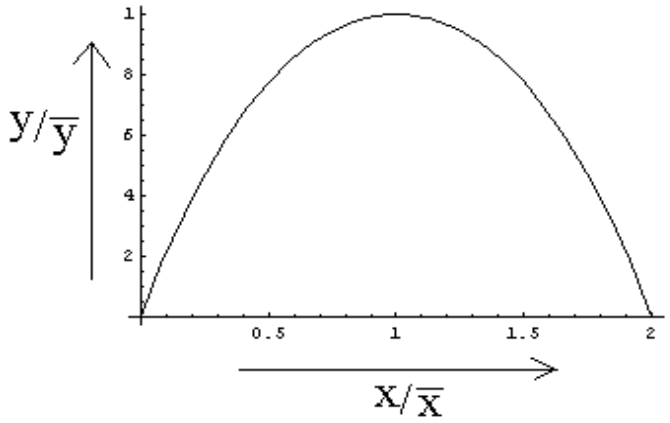


Figure 1.3: Light-path solution for a linear nonuniform medium

1.2 Geometric Formulation of Ray Optics

1.2.1 Frenet-Serret Curvature of Light Path

We now return to the general formulation for light-ray propagation based on the time integral (1.1), where the integrand is

$$F\left(\mathbf{x},\, \frac{d\mathbf{x}}{d\sigma}\right) \;=\; n(\mathbf{x}) \;\left|\frac{d\mathbf{x}}{d\sigma}\right|,$$

where light rays are now allowed to travel in three-dimensional space and the index of refraction $n(\mathbf{x})$ is a general function of position. Euler's First equation in this case is

$$\frac{d}{d\sigma}\left(\frac{\partial F}{\partial(d\mathbf{x}/d\sigma)}\right) \;=\; \frac{\partial F}{\partial \mathbf{x}}, \tag{1.18}$$

where

$$\frac{\partial F}{\partial(d\mathbf{x}/d\sigma)} \;=\; \frac{n}{\lambda}\,\frac{d\mathbf{x}}{d\sigma} \quad \text{and} \quad \frac{\partial F}{\partial \mathbf{x}} \;=\; \lambda\,\nabla n,$$

with

$$\lambda \;=\; \left|\frac{d\mathbf{x}}{d\sigma}\right| \;=\; \sqrt{\left(\frac{dx}{d\sigma}\right)^2 \;+\; \left(\frac{dy}{d\sigma}\right)^2 \;+\; \left(\frac{dz}{d\sigma}\right)^2}.$$

Euler's First Equation (1.18), therefore, becomes

$$\frac{d}{d\sigma}\left(\frac{n}{\lambda}\,\frac{d\mathbf{x}}{d\sigma}\right) \;=\; \lambda\,\nabla n. \tag{1.19}$$

Euler's Second Equation, on the other hand, states that

$$F \; - \; \frac{d\mathbf{x}}{d\sigma} \cdot \frac{\partial F}{\partial (d\mathbf{x}/d\sigma)} \; = \; 0$$

is a constant of motion.

By choosing the ray parametrization $d\sigma = ds$ (so that $\lambda = 1$), we find that the ray *velocity* $d\mathbf{x}/ds = \widehat{\mathsf{k}}$ is a unit vector which defines the direction of the wave vector $\mathbf{k}$. With this parametrization, Euler's equation (1.19) is now replaced with

$$\frac{d}{ds}\left(n \, \frac{d\mathbf{x}}{ds} \right) \; = \; \nabla n \quad \rightarrow \quad \frac{d^2\mathbf{x}}{ds^2} \; = \; \frac{d\mathbf{x}}{ds} \times \left(\nabla \ln n \times \frac{d\mathbf{x}}{ds} \right). \tag{1.20}$$

Eq. (1.20) shows that the *Frenet-Serret* curvature of the light path is $|\nabla \ln n \times \widehat{\mathsf{k}}\,|$ (and its radius of curvature is $|\nabla \ln n \times \widehat{\mathsf{k}}\,|^{-1}$) while the path has zero torsion since it is planar.

Eq. (1.20) can also be written in geometric form by introducing the metric relation $ds^2 \; = \; g_{ij}\, dx^i\, dx^j$, where $g_{ij} \; = \; \mathsf{e}_i \cdot \mathsf{e}_j$ denotes the metric tensor defined in terms of the contravariant-basis vectors $(\mathsf{e}_1, \mathsf{e}_2, \mathsf{e}_3)$, so that

$$\frac{d\mathbf{x}}{ds} \; = \; \frac{dx^i}{ds} \, \mathsf{e}_i .$$

Using the definition for the Christoffel symbol

$$\Gamma^{\ell}_{jk} \; = \; \frac{1}{2} \, g^{\ell i} \left(\frac{\partial g_{ij}}{\partial x^k} \; + \; \frac{\partial g_{ik}}{\partial x^j} \; - \; \frac{\partial g_{jk}}{\partial x^i} \right),$$

where g^{ij} denotes a component of the inverse metric (i.e., $g^{ij}\, g_{jk} = \delta^i{}_k$), we find the relations

$$\frac{d\mathsf{e}_j}{ds} \; = \; \Gamma^i_{jk} \, \frac{dx^k}{ds} \, \mathsf{e}_i ,$$

and

$$\frac{d^2\mathbf{x}}{ds^2} \; = \; \frac{d^2x^i}{ds^2} \, \mathsf{e}_i \; + \; \frac{dx^i}{ds} \, \frac{d\mathsf{e}_i}{ds} \; = \; \left(\frac{d^2x^i}{ds^2} \; + \; \Gamma^i_{jk} \, \frac{dx^j}{ds} \, \frac{dx^k}{ds} \right) \mathsf{e}_i .$$

By combining these relations, Eq. (1.20) becomes

$$\frac{d^2x^i}{ds^2} \; + \; \Gamma^i_{jk} \, \frac{dx^j}{ds} \, \frac{dx^k}{ds} \; = \; \left(g^{ij} \; - \; \frac{dx^i}{ds} \, \frac{dx^j}{ds} \right) \frac{\partial \ln n}{\partial x^j}, \tag{1.21}$$

Looking at Figure 1.4, we see that the light ray bends towards regions of higher index of refraction. Note that, if we introduce the unit vector $\widehat{\mathsf{n}} = \nabla n/(|\nabla n|)$ pointing in the direction of increasing index of refraction, we find the equation

$$\frac{d}{ds}\left(\widehat{\mathsf{n}} \times n \, \frac{d\mathbf{x}}{ds} \right) \; = \; \frac{d}{ds}\left(\widehat{\mathsf{n}} \times n \, \widehat{\mathsf{k}} \right) \; = \; \frac{d\widehat{\mathsf{n}}}{ds} \times n \, \widehat{\mathsf{k}},$$

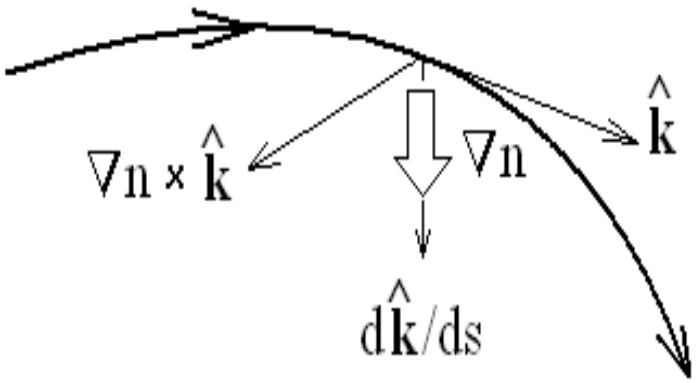


Figure 1.4: Light curvature

where we have used Eq. (1.20) to obtain

$$\hat{\mathsf{n}} \times \frac{d}{ds}\left(n\,\frac{d\mathbf{x}}{ds}\right) \;=\; \hat{\mathsf{n}} \times \nabla n \;=\; 0.$$

Hence, if the direction $\hat{\mathsf{n}}$ is constant along the path of a light ray (i.e., $d\hat{\mathsf{n}}/ds = 0$), then the quantity $\hat{\mathsf{n}} \times n\,\hat{\mathsf{k}}$ is a constant. In addition, when a light ray progagates in two dimensions, this conservation law implies that the quantity $|\hat{\mathsf{n}} \times n\,\hat{\mathsf{k}}| = n\,\sin\theta$ is also a constant, which is none other than Snell's Law (1.10).

1.2.2 Light Propagation in Spherical Geometry

By using the general ray-orbit equation (1.20), we can also show that for a spherically-symmetric nonuniform medium with index of refraction $n(r)$, the light-ray orbit $\mathbf{r}(s)$ satisfies the conservation law

$$\frac{d}{ds}\left(\mathbf{r} \times n(r)\,\frac{d\mathbf{r}}{ds}\right) \;=\; 0. \tag{1.22}$$

Here, we use the fact that the ray-orbit path is planar and, thus, we write

$$\mathbf{r} \times \frac{d\mathbf{r}}{ds} \;=\; r\,\sin\phi\,\hat{\mathsf{z}}, \tag{1.23}$$

where ϕ denotes the angle between the position vector $\mathbf{r}$ and the tangent vector $d\mathbf{r}/ds$ (see Figure 1.5). The conservation law (1.22) for ray orbits in a spherically-symmetric medium

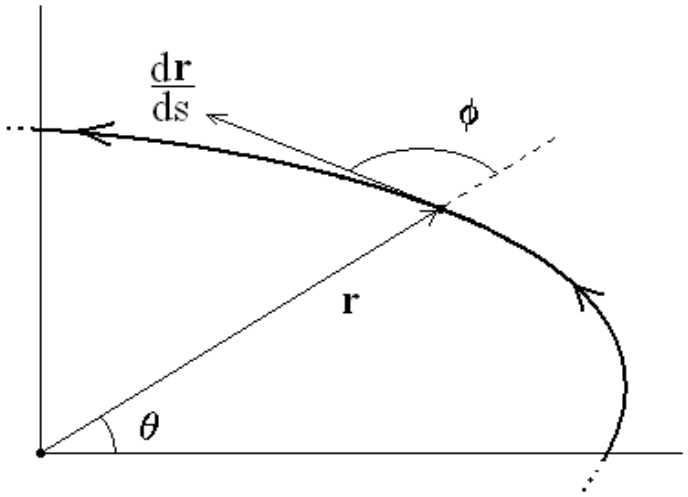


Figure 1.5: Light path in a nonuniform medium with spherical symmetry

can, therefore, be expressed as

$$n(r)\, r\ \sin\phi(r)\ =\ N\, a,$$

where N and a are constants (determined from initial conditions); note that the condition $n(r)\, r > N\, a$ must be satisfied.

An explicit expression for the ray orbit $r(\theta)$ is obtained as follows. First, since $d\mathbf{r}/ds$ is a unit vector, we find

$$\frac{d\mathbf{r}}{ds}\ =\ \frac{d\theta}{ds}\left(r\,\widehat{\theta}\ +\ \frac{dr}{d\theta}\,\widehat{r}\right)\ =\ \frac{r\,\widehat{\theta}\ +\ (dr/d\theta)\,\widehat{r}}{\sqrt{r^2\ +\ (dr/d\theta)^2}},$$

so that

$$\frac{d\theta}{ds}\ =\ \frac{1}{\sqrt{r^2\ +\ (dr/d\theta)^2}}$$

and Eq. (1.23) yields

$$\mathbf{r}\times\frac{d\mathbf{r}}{ds}\ =\ r\ \sin\phi\,\widehat{\mathsf{z}}\ =\ r^2\,\frac{d\theta}{ds}\,\widehat{\mathsf{z}}\quad\rightarrow\quad\sin\phi\ =\ \frac{r}{\sqrt{r^2\ +\ (dr/d\theta)^2}}\ =\ \frac{Na}{nr}.$$

Next, integration by quadrature yields

$$\frac{dr}{d\theta}\ =\ \frac{r}{Na}\,\sqrt{n(r)^2\, r^2 - N^2\, a^2}\quad\rightarrow\quad\theta(r)\ =\ N\, a\,\int_{r_0}^{r}\frac{d\rho}{\rho\,\sqrt{n^2(\rho)\,\rho^2 - N^2\, a^2}},$$

where $\theta(r_0) = 0$. Lastly, a change of integration variable $\eta = Na/\rho$ yields

$$\theta(r) \;=\; \int_{Na/r}^{Na/r_0} \frac{d\eta}{\sqrt{\overline{n}^2(\eta) - \eta^2}}, \tag{1.24}$$

where $\overline{n}(\eta) \equiv n(Na/\eta)$.

Consider, for example, the spherically-symmetric refractive index

$$n(r) \;=\; n_0 \sqrt{2 \;-\; \frac{r^2}{R^2}} \quad\rightarrow\quad \overline{n}^2(\eta) \;=\; n_0^2 \left(2 \;-\; \frac{N^2\epsilon^2}{\eta^2}\right),$$

where $n_0 = n(R)$ denotes the refractive index at $r = R$ and $\epsilon = a/R$ is a dimensionless parameter. Hence, Eq. (1.24) becomes

$$\theta(r) \;=\; \int_{Na/r}^{Na/r_0} \frac{\eta\, d\eta}{\sqrt{2n_0^2\,\eta^2 - n_0^2 N^2\epsilon^2 - \eta^4}} \;=\; \frac{1}{2} \int_{(Na/r)^2}^{(Na/r_0)^2} \frac{d\sigma}{\sqrt{n_0^4\,\mathsf{e}^2 - (\sigma - n_0^2)^2}},$$

where $\mathsf{e} = \sqrt{1 - N^2\epsilon^2/n_0^2}$ (assuming that $n_0 > N\,\epsilon$). Using the trigonometric substitution $\sigma = n_0^2\,(1 + \mathsf{e}\,\cos\chi)$, we find

$$\theta(r) \;=\; \frac{1}{2}\,\chi(r) \quad\rightarrow\quad r^2(\theta) \;=\; \frac{r_0^2\,(1+\mathsf{e})}{1 \;+\; \mathsf{e}\,\cos 2\theta},$$

which represents an ellipse of major radius and minor radius

$$r_1 \;=\; R\,(1+\mathsf{e})^{1/2} \quad\text{and}\quad r_0 \;=\; R\,(1-\mathsf{e})^{1/2}$$

respectively. The angle $\phi(\theta)$ defined from the conservation law (1.22) is now expressed as

$$\sin\phi(\theta) \;=\; \frac{1 \;+\; \mathsf{e}\,\cos 2\theta}{\sqrt{1 + \mathsf{e}^2 + 2\,\mathsf{e}\,\cos 2\theta}},$$

so that $\phi = \frac{\pi}{2}$ at $\theta = 0$ and $\frac{\pi}{2}$, as expected for an ellipse.

1.2.3 Geodesic Representation of Light Propagation

We now investigate the geodesic properties of light propagation in a nonuniform refractive medium. For this purpose, let us consider a path AB in space from point A to point B parametrized by the continuous parameter σ, i.e., $\mathbf{x}(\sigma)$ such that $\mathbf{x}(A) = \mathbf{x}_A$ and $\mathbf{x}(B) = \mathbf{x}_B$. The time taken by light in propagating from A to B is

$$T_{AB} \;=\; \int_A^B \frac{dt}{d\sigma}\, d\sigma \;=\; \int_A^B \frac{n}{c} \left(g_{ij}\,\frac{dx^i}{d\sigma}\,\frac{dx^j}{d\sigma}\right)^{1/2} d\sigma, \tag{1.25}$$

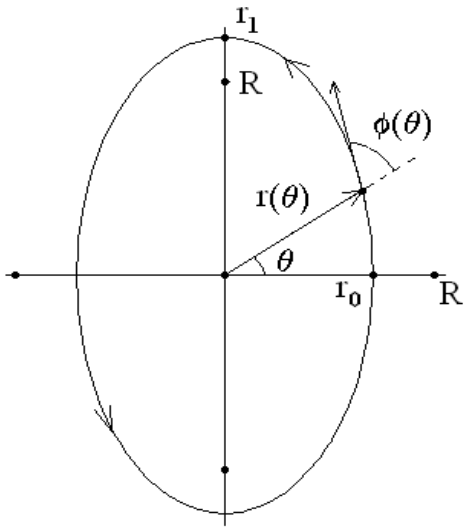


Figure 1.6: Light elliptical path

where $dt = n\,ds/c$ denotes the infinitesimal time interval taken by light in moving an infinitesimal distance ds in a medium with refractive index n and the space metric is denoted by g_{ij}.

We now define the medium-modified space metric $\overline{g}_{ij} = n^2\, g_{ij}$ as

$$c^2\, dt^2 \;=\; n^2\; ds^2 \;=\; n^2\, g_{ij}\; dx^i\, dx^j \;=\; \overline{g}_{ij}\, dx^i\, dx^j,$$

and apply the Principle of Least Time by considering geodesic motion associated with the medium-modified space metric $\overline{g}_{ij}$. The variation in time δT_{AB} is given (to first order in δx^i) as

$$\begin{aligned}
\delta T_{AB} &= \frac{1}{2c^2}\int_A^B \frac{d\sigma}{dt/d\sigma}\left[\frac{\partial \overline{g}_{ij}}{\partial x^k}\,\delta x^k\,\frac{dx^i}{d\sigma}\,\frac{dx^j}{d\sigma} \;+\; 2\,\overline{g}_{ij}\,\frac{d\delta x^i}{d\sigma}\,\frac{dx^j}{d\sigma}\right] \\
&= \frac{1}{2c^2}\int_{t_A}^{t_B}\left[\frac{\partial \overline{g}_{ij}}{\partial x^k}\,\delta x^k\,\frac{dx^i}{dt}\,\frac{dx^j}{dt} \;+\; 2\,\overline{g}_{ij}\,\frac{d\delta x^i}{dt}\,\frac{dx^j}{dt}\right] dt.
\end{aligned}$$

By integrating by parts the second term we obtain

$$\begin{aligned}
\delta T_{AB} &= -\int_{t_A}^{t_B}\left[\frac{d}{dt}\left(\overline{g}_{ij}\,\frac{dx^j}{dt}\right) \;-\; \frac{1}{2}\,\frac{\partial \overline{g}_{jk}}{\partial x^i}\,\frac{dx^j}{dt}\,\frac{dx^k}{dt}\right]\delta x^i\,\frac{dt}{c^2} \\
&= -\int_{t_A}^{t_B}\left[\overline{g}_{ij}\,\frac{d^2x^j}{dt^2} \;+\; \left(\frac{\partial \overline{g}_{ij}}{\partial x^k} \;-\; \frac{1}{2}\,\frac{\partial \overline{g}_{jk}}{\partial x^i}\right)\frac{dx^j}{dt}\,\frac{dx^k}{dt}\right]\delta x^i\,\frac{dt}{c^2}.
\end{aligned}$$

We now note that the second term can be written as

$$\left(\frac{\partial \overline{g}_{ij}}{\partial x^k} \;-\; \frac{1}{2}\,\frac{\partial \overline{g}_{jk}}{\partial x^i}\right)\frac{dx^j}{dt}\,\frac{dx^k}{dt} \;=\; \frac{1}{2}\left(\frac{\partial \overline{g}_{ij}}{\partial x^k} \;+\; \frac{\partial \overline{g}_{ik}}{\partial x^j} \;-\; \frac{\partial \overline{g}_{jk}}{\partial x^i}\right)\frac{dx^j}{dt}\,\frac{dx^k}{dt}$$

$$= \overline{\Gamma}_{i|jk} \frac{dx^j}{dt} \frac{dx^k}{dt},$$

using the definition of the Christoffel symbol

$$\overline{\Gamma}^{\ell}_{jk} = \overline{g}^{\ell i} \overline{\Gamma}_{i|jk} = \frac{\overline{g}^{\ell i}}{2} \left(\frac{\partial \overline{g}_{ij}}{\partial x^k} + \frac{\partial \overline{g}_{ik}}{\partial x^j} - \frac{\partial \overline{g}_{jk}}{\partial x^i} \right),$$

where $\overline{g}^{ij} = n^{-2} g^{ij}$ denotes components of the inverse metric (i.e., $\overline{g}^{ij} \overline{g}_{jk} = \delta^i{}_k$), and its symmetry property with respect to interchange of its two covariant indices ($j \leftrightarrow k$). Hence, the variation δT_{AB} can be expressed as

$$\delta T_{AB} = \int_{t_A}^{t_B} \left[\frac{d^2 x^i}{dt^2} + \overline{\Gamma}^{i}_{jk} \frac{dx^j}{dt} \frac{dx^k}{dt} \right] \overline{g}_{i\ell} \, \delta x^{\ell} \frac{dt}{c^2}. \tag{1.26}$$

We, therefore, find that the light path $\mathbf{x}(t)$ is a solution of the geodesic equation

$$\frac{d^2 x^i}{dt^2} + \overline{\Gamma}^{i}_{jk} \frac{dx^j}{dt} \frac{dx^k}{dt} = 0, \tag{1.27}$$

which is also the path of least time for which $\delta T_{AB} = 0$.

1.2.4 Eikonal Representation

Lastly, the index of refraction itself (for an isotropic medium) can be written as

$$n = |\nabla \mathcal{S}| = \frac{ck}{\omega} \quad \text{or} \quad \nabla \mathcal{S} = n \frac{d\mathbf{x}}{ds} = \frac{c\,\mathbf{k}}{\omega}, \tag{1.28}$$

where $\mathcal{S}$ is called the *eikonal* function and the phase speed of a light wave is ω/k; note that the surface $\mathcal{S}(x, y, z) = constant$ represents a *wave-front*, which is complementary to the ray picture used so far. To show that this definition is consistent with Eq. (1.20), we easily check that

$$\frac{d}{ds} \left(n \frac{d\mathbf{x}}{ds} \right) = \frac{d\nabla \mathcal{S}}{ds} = \frac{d\mathbf{x}}{ds} \cdot \nabla \nabla \mathcal{S} = \frac{\nabla \mathcal{S}}{n} \cdot \nabla \nabla \mathcal{S} = \frac{1}{2n} \nabla |\nabla \mathcal{S}|^2 = \nabla n.$$

This definition, therefore, implies that $\nabla \times \mathbf{k} = 0$ (since $\nabla \times \nabla \mathcal{S} = 0$), where we used the fact that the frequency of a wave is unchanged by refraction.

1.3 Brachistochrone Problem

The brachistochrone problem is another least-time problem and was first solved in 1696 by Johann Bernoulli (1667-1748). The problem can be stated as follows. A bead is released

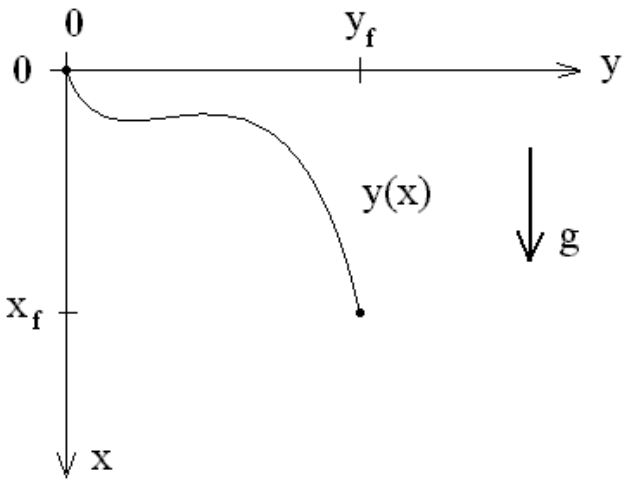


Figure 1.7: Brachistrochrone problem

from rest and slides down a frictionless wire that connects the origin to a given point (x_f, y_f); see Figure 1.7. The question posed by the brachistochrone problem is to determine the shape $y(x)$ of the wire for which the descent of the bead under gravity takes the shortest amount of time. Using the (x, y)-coordinates shown above, the speed of the bead after it has fallen a vertical distance x along the wire is $v = \sqrt{2g\,x}$ (where g denotes the gravitational acceleration) and, thus, the time integral

$$T[y] \;=\; \int_0^{x_f} \frac{\sqrt{1+(y')^2}}{\sqrt{2\,gx}}\,dx \;=\; \int_0^{x_f} F(y, y', x)\,dx, \tag{1.29}$$

is a functional of the path $y(x)$. Since the integrand of Eq. (1.29) is independent of the y-coordinate ($\partial F/\partial y = 0$), the Euler's First Equation (1.5) simply yields

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) \;=\; 0 \quad\rightarrow\quad \frac{\partial F}{\partial y'} \;=\; \frac{y'}{\sqrt{2\,gx\;[1+(y')^2]}} \;=\; \alpha,$$

where α is a constant, which leads to

$$\frac{(y')^2}{1+(y')^2} \;=\; \frac{x}{a},$$

where $a = (2\alpha^2 g)^{-1}$ is a scale length for the problem. Integration by quadrature yields the integral solution

$$y(x) \;=\; \int_0^x \sqrt{\frac{\eta}{a-\eta}}\,d\eta,$$

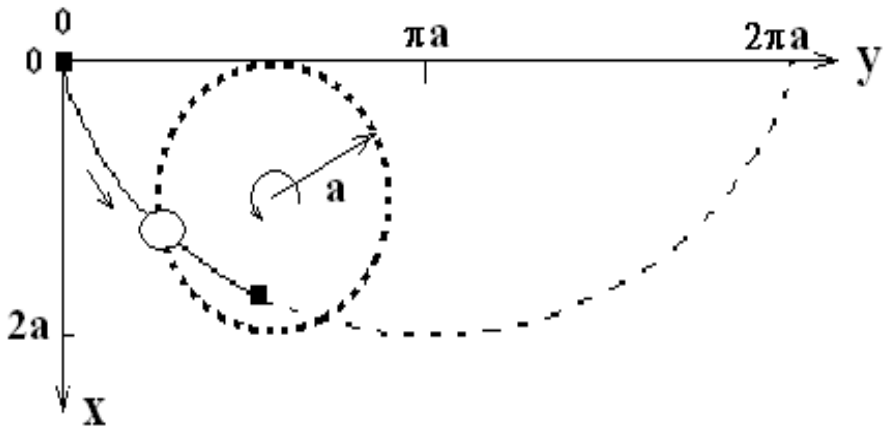


Figure 1.8: Brachistochrone solution

subject to the initial condition $y(x = 0) = 0$. Using the trigonometric substitution

$$\eta \;=\; 2a\,\sin^2(\theta/2) \;=\; a\,(1 \;-\; \cos\theta),$$

we obtain the parametric solution

$$x(\theta) \;=\; a\,(1 \;-\; \cos\theta) \quad \text{and} \quad y(\theta) \;=\; a\,(\theta \;-\; \sin\theta). \tag{1.30}$$

This solution yields a parametric representation of the *cycloid* (Figure 1.8) where the bead is placed on a rolling hoop of radius a. Lastly, the time integral (1.29) for the cycloid solution (1.30) is

$$T_{cycloid}(\Theta) \;=\; \int_0^{\Theta} \frac{\sqrt{(dx/d\theta)^2 + (dy/d\theta)^2}\; d\theta}{\sqrt{2ga\,(1-\cos\theta)}} \;=\; \Theta\,\sqrt{\frac{a}{g}}.$$

In particular, the time needed to reach the bottom of the cycloid ($\Theta = \pi$) is $\pi\,\sqrt{a/g}$.

1.4 Problems

Problem 1

According to the Calculus of Variations, the straight line $y(x) = mx$ between the two points $(0, 0)$ and $(1, m)$ on the (x, y)-plane yields a minimum value for the length integral

$$L[y] \;= \int_0^1 \sqrt{1 + (y')^2}\, dx,$$

since the path $y(x) = mx$ satisfies the Euler equation

$$0 \;=\; \frac{d}{dx}\left(\frac{y'}{\sqrt{1 + (y')^2}}\right) \;=\; \frac{y''}{[1 + (y')^2]^{3/2}}.$$

By choosing the path variation $\epsilon\,\delta y(x) = \epsilon\, x(x-1)$, which vanishes at $x = 0$ and 1, we find

$$L[y + \epsilon\,\delta y] \;=\; \int_0^1 \sqrt{1 \;+\; [2\,\epsilon\, x + (m - \epsilon)]^2}\, dx \;=\; \frac{1}{2\epsilon} \int_{\tan^{-1}(m-\epsilon)}^{\tan^{-1}(m+\epsilon)} \sec^3\theta\; d\theta.$$

Evaluate $L[y + \epsilon\,\delta y]$ explicitly as a function of m and ϵ and show that it has a minimum in ϵ at $\epsilon = 0$.

Problem 2

Prove the identities (1.16).

Problem 3

A light ray travels in a medium with refractive index

$$n(y) \;=\; n_0\, \exp(-\beta\, y),$$

where n_0 is the refractive index at $y = 0$ and β is a positive constant.

(a) Use the results of the Principle of Least Time contained in the Notes distributed in class to show that the path of the light ray is expressed as

$$y(x;\beta) \;=\; \frac{1}{\beta}\, \ln\left[\, \frac{\cos(\beta\, x - \phi_0)}{\cos\phi_0}\, \right], \tag{1.31}$$

where the light ray is initially travelling upwards from $(x, y) = (0, 0)$ at an angle ϕ_0.

(b) Using the appropriate mathematical techniques, show that we recover the expected result

$$\lim_{\beta\to 0}\; y(x;\beta) \;=\; (\tan\phi_0)\; x$$

from Eq. (1.31).

(c) The light ray reaches a maximum height $\overline{y}$ at $x = \overline{x}(\beta)$, where $y'(\overline{x};\beta) = 0$. Find expressions for $\overline{x}$ and $\overline{y}(\beta) = y(\overline{x};\beta)$.

Chapter 2

Lagrangian Mechanics

2.1 Maupertuis-Jacobi Principle of Least Action

The publication of Fermat's Principle of Least Time in 1657 generated an intense controversy between Fermat and disciples of René Descartes (1596-1650) involving whether light travels slower (Fermat) or faster (Descartes) in a dense medium as compared to free space.

In 1740, Pierre Louis Moreau de Maupertuis (1698-1759) stated (without proof) that, in analogy with Fermat's Principle of Least Time for light, a particle of mass m under the influence of a force $\mathbf{F}$ moves along a path which satisfies the Principle of *Least Action*: $\delta S = 0$, where the action integral is defined as

$$S \;=\; \int \mathbf{p} \cdot d\mathbf{x} \;=\; \int mv\, ds, \tag{2.1}$$

where $v = ds/dt$ denotes the magnitude of particle velocity, which can also be expressed as $v = \sqrt{\frac{2}{m}(E-U)}$, with the particle's kinetic energy K written in terms of its total energy E and its potential energy U.

In 1744, Euler proved the Principle of Least Action $\delta \int mv\, ds = 0$ for particle motion in the (x, y)-plane as follows. For this purpose, we use the Frenet-Serret curvature formula for the path $y(x)$; here, we define the tangent unit vector $\widehat{\mathsf{v}}$ and the principal normal unit vector $\widehat{\mathsf{n}}$ as

$$\widehat{\mathsf{v}} \;=\; \frac{d\mathbf{x}}{ds} \;=\; \frac{\widehat{\mathsf{x}} \;+\; y'\,\widehat{\mathsf{y}}}{\sqrt{1+(y')^2}} \quad \text{and} \quad \widehat{\mathsf{n}} \;=\; \frac{y'\,\widehat{\mathsf{x}} \;-\; \widehat{\mathsf{y}}}{\sqrt{1+(y')^2}}, \tag{2.2}$$

where $y' = dy/dx$ and $ds = dx\,\sqrt{1+(y')^2}$. The Frenet-Serret formula for the curvature of a two-dimensional curve is

$$\frac{d\widehat{\mathsf{v}}}{ds} \;=\; \frac{|y''|\;\widehat{\mathsf{n}}}{[1+(y')^2]^{3/2}} \;=\; \kappa\,\widehat{\mathsf{n}},$$

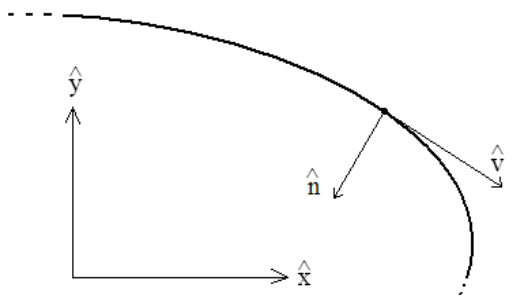


Figure 2.1: Frenet-Serret frame

where the instantaneous radius of curvature ρ is defined as $\rho = \kappa^{-1}$ (see Figure 2.1). First, by using Newton's Second Law of Motion and the Energy conservation law, we find the relation

$$\mathbf{F} \;=\; mv\left(\frac{dv}{ds}\,\hat{\mathsf{v}} \;+\; v\,\frac{d\hat{\mathsf{v}}}{ds}\right) \;=\; \hat{\mathsf{v}}\,(\hat{\mathsf{v}}\cdot\nabla K) \;+\; mv^2\,\kappa\,\hat{\mathsf{n}} \;=\; \nabla K \tag{2.3}$$

between the unit vectors $\hat{\mathsf{v}}$ and $\hat{\mathsf{n}}$ associated with the path, the Frenet-Serret curvature κ, and the kinetic energy $K = \frac{1}{2}\,mv^2(x,y)$ of the particle. Note that Eq. (2.3) can be re-written as

$$\frac{d\hat{\mathsf{v}}}{ds} \;=\; \hat{\mathsf{v}}\times\,(\nabla \ln v \times \hat{\mathsf{v}})\,, \tag{2.4}$$

which hightlights a deep connection with Eq. (1.20) derived from Fermat's Principle of Least Time, where the index of refraction n is now replaced by $v = \sqrt{\frac{2}{m}(E-U)}$. Lastly, we point out that the type of dissipationless forces considered in Eq. (2.3) involves *active* forces (defined as forces that do work), as opposed to *passive* forces (defined as forces that do no work, such as constraint forces).

Next, the action integral (2.1) is expressed as

$$S \;=\; \int\, m\,v(x,y)\,\sqrt{1+(y')^2}\,dx \;=\; \int\, F(y,y';x)\,dx, \tag{2.5}$$

so that the Euler's First Equation (1.5) corresponding to Maupertuis' action integral (2.5), with

$$\frac{\partial F}{\partial y'} \;=\; \frac{mv\,y'}{\sqrt{1+(y')^2}} \quad\text{and}\quad \frac{\partial F}{\partial y} \;=\; m\,\sqrt{1+(y')^2}\,\frac{\partial v}{\partial y},$$

yields the Maupertuis-Euler equation

$$\frac{m\,y'}{\sqrt{1+(y')^2}}\,\frac{\partial v}{\partial x} \;+\; \frac{m\,y''}{[1+(y')^2]^{3/2}} \;=\; m\,\sqrt{1+(y')^2}\,\frac{\partial v}{\partial y}. \tag{2.6}$$

Using the relation $\mathbf{F} = \nabla K$ and the Frenet-Serret formulas (2.2), the Maupertuis-Euler equation (2.6) becomes

$$mv^2\,\kappa \;=\; \mathbf{F}\cdot\widehat{\mathsf{n}},$$

from which we recover Newton's Second Law (2.3).

Carl Gustav Jacobi (1804-1851) emphasized the connection between Fermat's Principle of Least Time (1.1) and Maupertuis' Principle of Least Action (2.1) by introducing a different form of the Principle of Least Action $\delta S = 0$, where Jacobi's action integral is

$$S \;=\; \int\,\sqrt{2m\,(E-U)}\,ds \;=\; 2\,\int K\,dt, \tag{2.7}$$

where particle momentum is written as $p = \sqrt{2m\,(E-U)}$. To obtain the second expression of Jacobi's action integral (2.7), Jacobi made use of the fact that, by introducing a path parameter τ such that $v = ds/dt = s'/t'$ (where a prime here denotes a τ-derivative), we find

$$K \;=\; \frac{m\,(s')^2}{2\,(t')^2} \;=\; E \;-\; U,$$

so that

$$2\,K\,t' \;=\; s'\,\sqrt{2m\,(E-U)},$$

and the second form of Jacobi's action integral results. Jacobi used the Principle of Least Action (2.7) to establish the geometric foundations of particle mechanics. Here, the Euler-Jacobi equation resulting from Jacobi's Principle of Least Action is expressed as

$$\frac{d}{ds}\left(\sqrt{E-U}\,\frac{d\mathbf{x}}{ds}\right) \;=\; \nabla\sqrt{E-U}\,,$$

which is identical in form to the light-curvature equation (1.20).

2.2 Principle of Least Action of Euler and Lagrange

2.2.1 Generalized Coordinates in Configuration Space

The configuration of a mechanical system with constraints evolving in n-dimensional space, with spatial coordinates $\mathbf{x} = (x^1, x^2, ..., x^n)$, can sometimes be described in terms of *generalized* coordinates $\mathbf{q} = (q^1, q^2, ..., q^k)$ in a k-dimensional *configuration* space, with $k \leq n$. For example, for a mechanical system composed of two particles (see Figure 2.2), with masses (m_1, m_2) and three-dimensional coordinate positions $(\mathbf{x}_1, \mathbf{x}_2)$, tied together with a massless rod (so that the distance $|\mathbf{x}_1 - \mathbf{x}_2|$ is constant), the configuration of this two-particle system can be described in terms of the coordinates $\mathbf{x}_{CM} = (m_1\,\mathbf{x}_1 + m_2\,\mathbf{x}_2)/(m_1 + m_2)$ of the center-of-mass (CM) in the Laboratory frame (O) and the orientation of the rod in the CM frame (O') expressed in terms of the two angles (θ, φ). Hence, as a result of the

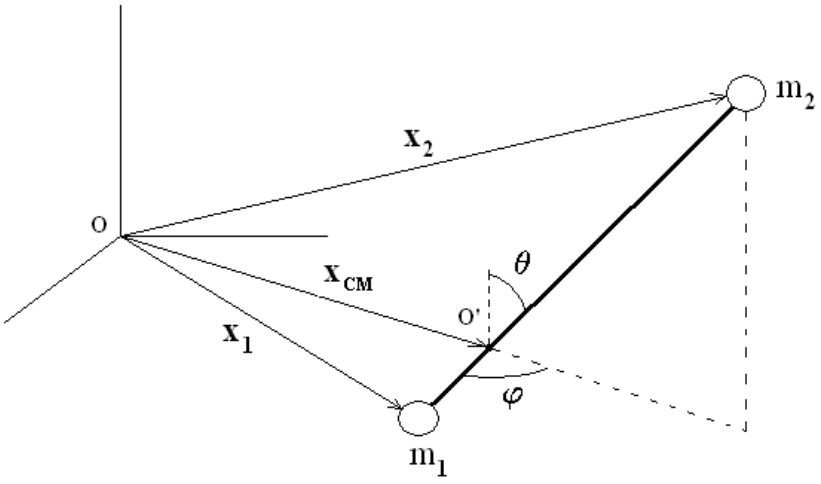


Figure 2.2: Configuration space

existence of a single constraint, the generalized coordinates for this system are $(\mathbf{x}_{CM};\, \theta, \varphi)$ and we have reduced the number of coordinates needed to describe the state of the system from six to five.

2.2.2 Constrained Motion on a Surface

The general problem associated with the motion of a particle constrained to move on a surface described by the relation $F(x, y, z) = 0$ is described as follows. First, since the velocity $d\mathbf{x}/dt$ of the particle along its trajectory must be perpendicular to the gradient ∇F, we find that $d\mathbf{x} \cdot \nabla F = 0$. Next, any point $\mathbf{x}$ on the surface $F(x, y, z) = 0$ may be parametrized by two *surface* coordinates (u, v) such that

$$\frac{\partial \mathbf{x}}{\partial u}(u, v) \cdot \nabla F \;=\; 0 \;=\; \frac{\partial \mathbf{x}}{\partial v}(u, v) \cdot \nabla F.$$

Hence, we may write

$$d\mathbf{x} \;=\; \frac{\partial \mathbf{x}}{\partial u}\, du \;+\; \frac{\partial \mathbf{x}}{\partial v}\, dv \quad \text{and} \quad \frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} \;=\; \mathcal{J}\, \nabla F,$$

where the function $\mathcal{J}$ depends on the surface coordinates (u, v). It is, thus, quite clear that the surface coordinates (u, v) are the generalized coordinates for this constrained motion.

For example, we consider the motion of a particle constrained to move on the surface of a cone of apex angle α. Here, the constraint is expressed as $F(x, y, z) = r - z \tan\alpha = 0$

with $\nabla F = \widehat{r} - \tan\alpha\, \widehat{z}$. The surface coordinates can be chosen to be the polar angle θ and the function

$$s(x,y,z) \;=\; \sqrt{x^2+y^2+z^2},$$

which measures the distance from the apex of the cone (defining the origin), so that

$$\frac{\partial \mathbf{x}}{\partial \theta} \;=\; r\,\widehat{\theta} \;=\; r\,\widehat{z}\times\widehat{r} \quad \text{and} \quad \frac{\partial \mathbf{x}}{\partial s} \;=\; \sin\alpha\,\widehat{r} \;+\; \cos\alpha\,\widehat{z} \;=\; \widehat{s},$$

with

$$\frac{\partial \mathbf{x}}{\partial \theta}\times\frac{\partial \mathbf{x}}{\partial s} \;=\; r\,\cos\alpha\,\nabla F$$

and $\mathcal{J} = r\,\cos\alpha$. Lastly, the velocity of the particle is $\dot{\mathbf{x}} = \dot{s}\,\widehat{s} + r\,\dot{\theta}\,\widehat{\theta}$ and, thus, it satisfies $\dot{\mathbf{x}}\cdot\nabla F = 0$. We shall return to this example in Sec. 2.4.4.

2.2.3 Euler-Lagrange Equations

The Principle of Least Action (also known as Hamilton's principle as it is formulated here) is expressed in terms of a function $L(\mathbf{q},\dot{\mathbf{q}};\,t)$ known as the *Lagrangian*, which appears in the *action* integral

$$S[\mathbf{q}] \;=\; \int_{t_i}^{t_f} L(\mathbf{q},\dot{\mathbf{q}};\,t)\,dt, \tag{2.8}$$

where the action integral is a functional of the vector function $\mathbf{q}(t)$, which provides a path from the initial point $\mathbf{q}_i = \mathbf{q}(t_i)$ to the final point $\mathbf{q}_f = \mathbf{q}(t_f)$. The variational principle

$$0 \;=\; \delta S[\mathbf{q}] \;=\; \left(\frac{d}{d\epsilon}S[\mathbf{q}+\epsilon\,\delta\mathbf{q}]\right)_{\epsilon=0} \;=\; \int_{t_i}^{t_f}\delta\mathbf{q}\cdot\left[\;\frac{\partial L}{\partial \mathbf{q}} \;-\; \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right)\;\right]\,dt,$$

where the variation $\delta\mathbf{q}$ is assumed to vanish at the integration boundaries ($\delta\mathbf{q}_i = 0 = \delta\mathbf{q}_f$), yields the *Euler-Lagrange* equation for the generalized coordinate q^j $(j = 1, ..., k)$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^j}\right) \;=\; \frac{\partial L}{\partial q^j}. \tag{2.9}$$

The Lagrangian also satisfies the second Euler equation

$$\frac{d}{dt}\left(L \;-\; \dot{\mathbf{q}}\,\frac{\partial L}{\partial \dot{\mathbf{q}}}\right) \;=\; \frac{\partial L}{\partial t}, \tag{2.10}$$

and thus, for time-independent Lagrangian systems ($\partial L/\partial t = 0$), we find that $L - \dot{\mathbf{q}}\,\partial L/\partial\dot{\mathbf{q}}$ is a conserved quantity whose interpretation will be discussed shortly.

The form of the Lagrangian function $L(\mathbf{r},\dot{\mathbf{r}};\,t)$ is dictated by our requirement that Newton's Second Law $m\,\ddot{\mathbf{r}} = -\,\nabla U(\mathbf{r},t)$, which describes the motion of a particle of mass

m in a nonuniform (possibly time-dependent) potential $U(\mathbf{r}, t)$, be written in the Euler-Lagrange form (2.9). One easily obtains the form

$$L(\mathbf{r}, \dot{\mathbf{r}}; t) \;=\; \frac{m}{2}\,|\dot{\mathbf{r}}|^2 \;-\; U(\mathbf{r}, t), \tag{2.11}$$

for the Lagrangian of a particle of mass m, which is simply the kinetic energy of the particle **minus** its potential energy. The minus sign in Eq. (2.11) is important; not only does this form give us the correct equations of motion but, without the minus sign, energy would not be conserved. In fact, we note that Jacobi's action integral (2.7) can also be written as $A = \int [(K - U) + E]\, dt$, using the Energy conservation law $E = K + U$; hence, Energy conservation is the important connection between the Principles of Least Action of Maupertuis-Jacobi and Euler-Lagrange.

For a simple mechanical system, the Lagrangian function is obtained by computing the kinetic energy of the system and its potential energy and then construct Eq. (2.11). The construction of a Lagrangian function for a system of N particles proceeds in three steps as follows.

- **Step I.** Define k generalized coordinates $\{q^1(t), ..., q^k(t)\}$ that represent the instantaneous *configuration* of the mechanical system of N particles at time t.

- **Step II.** For each particle, construct the position vector $\mathbf{r}_a(\mathbf{q}; t)$ in Cartesian coordinates and its associated velocity

$$\mathbf{v}_a(\mathbf{q}, \dot{\mathbf{q}}; t) \;=\; \frac{\partial \mathbf{r}_a}{\partial t} \;+\; \sum_{j=1}^{k} \dot{q}^j\, \frac{\partial \mathbf{r}_a}{\partial q^j}$$

for $a = 1, ..., N$.

- **Step III.** Construct the kinetic energy

$$K(\mathbf{q}, \dot{\mathbf{q}}; t) \;=\; \sum_a \frac{m_a}{2}\,|\mathbf{v}_a(\mathbf{q}, \dot{\mathbf{q}}; t)|^2$$

and the potential energy

$$U(\mathbf{q}; t) \;=\; \sum_a U(\mathbf{r}_a(\mathbf{q}; t),\, t)$$

for the system and combine them to obtain the Lagrangian

$$L(\mathbf{q}, \dot{\mathbf{q}}; t) \;=\; K(\mathbf{q}, \dot{\mathbf{q}}; t) \;-\; U(\mathbf{q}; t).$$

From this Lagrangian $L(\mathbf{q}, \dot{\mathbf{q}}; t)$, the Euler-Lagrange equations (2.9) are derived for each generalized coordinate q^j:

$$\sum_a \frac{d}{dt}\left(\frac{\partial \mathbf{r}_a}{\partial q^j} \cdot m_a\, \mathbf{v}_a\right) \;=\; \sum_a \left(m_a\, \mathbf{v}_a \cdot \frac{\partial \mathbf{v}_a}{\partial q^j} \;-\; \frac{\partial \mathbf{r}_a}{\partial q^j} \cdot \nabla U\right), \tag{2.12}$$

where we have used the identity $\partial \mathbf{v}_a / \partial \dot{q}^j = \partial \mathbf{r}_a / \partial q^j$.

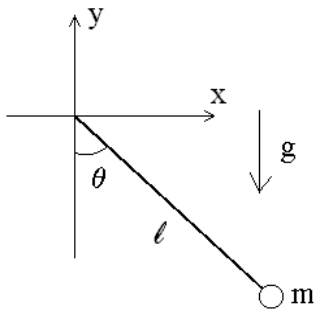


Figure 2.3: Generalized coordinates for the pendulum problem

2.3 Lagrangian Mechanics in Configuration Space

In this Section, we explore the Lagrangian formulation of several mechanical systems listed here in order of increasing complexity. As we proceed with our examples, we should realize how the Lagrangian formulation maintains its relative simplicity compared to the application of the more familiar Newton's method (Isaac Newton, 1643-1727) associated with the composition of forces.

2.3.1 Example I: Pendulum

As a first example, we consider a pendulum composed of an object of mass m and a massless string of constant length ℓ in a constant gravitational field with acceleration g. Although the motion of the pendulum is two-dimensional, a single generalized coordinate is needed to describe the configuration of the pendulum: the angle θ measured from the negative y-axis (see Figure 2.3). Here, the position of the object is given as

$$x(\theta) \;=\; \ell\,\sin\theta \quad \text{and} \quad y(\theta) \;=\; -\,\ell\,\cos\theta,$$

with associated velocity components

$$\dot{x}(\theta,\dot{\theta}) \;=\; \ell\,\dot{\theta}\,\cos\theta \quad \text{and} \quad \dot{y}(\theta,\dot{\theta}) \;=\; \ell\,\dot{\theta}\,\sin\theta.$$

Hence, the kinetic energy of the pendulum is

$$K \;=\; \frac{m}{2}\left(\dot{x}^2 \;+\; \dot{y}^2\right) \;=\; \frac{m}{2}\,\ell^2\dot{\theta}^2,$$

and choosing the zero potential energy point when $\theta = 0$ (see Figure above), the gravitational potential energy is

$$U \;=\; mg\ell\,(1-\cos\theta).$$

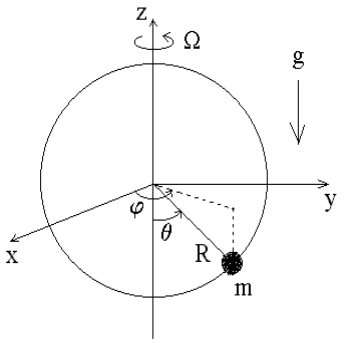


Figure 2.4: Generalized coordinates for the bead-on-a-rotating-hoop problem

The Lagrangian $L = K - U$ is, therefore, written as

$$L(\theta, \dot{\theta}) \;=\; \frac{m}{2}\,\ell^2\dot{\theta}^2 \;-\; mg\ell\,(1-\cos\theta),$$

and the Euler-Lagrange equation for θ is

$$\begin{aligned} \frac{\partial L}{\partial \dot{\theta}} \;=\; m\ell^2\,\dot{\theta} \;\;\rightarrow\;\; \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) &= m\ell^2\,\ddot{\theta} \\ \frac{\partial L}{\partial \theta} &= -\,mg\ell\,\sin\theta \end{aligned}$$

or

$$\ddot{\theta} \;+\; \frac{g}{\ell}\,\sin\theta \;=\; 0$$

The pendulum problem is solved in the next Chapter through the use of the Energy method (a simplified version of the Hamiltonian method). Note that, whereas the tension in the pendulum string must be considered explicitly in the Newtonian method, the string tension is replaced by the constraint $\ell =$ constant in the Lagrangian method.

2.3.2 Example II: Bead on a Rotating Hoop

As a second example, we consider a bead of mass m sliding freely on a hoop of radius R rotating with angular velocity Ω in a constant gravitational field with acceleration g. Here, since the bead of the rotating hoop moves on the surface of a sphere of radius R, we use the generalized coordinates given by the two angles θ (measured from the negative z-axis) and φ (measured from the positive x-axis), where $\dot{\varphi} = \Omega$. The position of the bead is given as

$$\begin{aligned} x(\theta,t) &= R\,\sin\theta\,\cos(\varphi_0+\Omega t), \\ y(\theta,t) &= R\,\sin\theta\,\sin(\varphi_0+\Omega t), \\ z(\theta,t) &= -\,R\,\cos\theta, \end{aligned}$$

where $\varphi(t) = \varphi_0 + \Omega\, t$, and its associated velocity components are

$$\begin{aligned}
\dot{x}(\theta, \dot{\theta}; t) &= R\left(\dot{\theta}\ \cos\theta\ \cos\varphi\ -\ \Omega\ \sin\theta\ \sin\varphi\right), \\
\dot{y}(\theta, \dot{\theta}; t) &= R\left(\dot{\theta}\ \cos\theta\ \sin\varphi\ +\ \Omega\ \sin\theta\ \cos\varphi\right), \\
\dot{z}(\theta, \dot{\theta}; t) &= R\,\dot{\theta}\ \sin\theta,
\end{aligned}$$

so that the kinetic energy of the bead is

$$K(\theta, \dot{\theta}) \ =\ \frac{m}{2}\,|\mathbf{v}|^2 \ =\ \frac{m\,R^2}{2}\left(\dot{\theta}^2\ +\ \Omega^2\ \sin^2\theta\right).$$

The gravitational potential energy is

$$U(\theta) \ =\ mgR\,(1\ -\ \cos\theta),$$

where we chose the zero potential energy point at $\theta = 0$ (see Figure 2.4). The Lagrangian $L = K - U$ is, therefore, written as

$$L(\theta, \dot{\theta}) \ =\ \frac{m\,R^2}{2}\left(\dot{\theta}^2\ +\ \Omega^2\ \sin^2\theta\right)\ -\ mgR\,(1 - \cos\theta),$$

and the Euler-Lagrange equation for θ is

$$\begin{aligned}
\frac{\partial L}{\partial \dot{\theta}} \ =\ mR^2\,\dot{\theta}\ \ \rightarrow\ \ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) &= mR^2\,\ddot{\theta} \\
\frac{\partial L}{\partial \theta} &= -\,mgR\,\sin\theta \\
&\quad +\ mR^2\Omega^2\ \cos\theta\ \sin\theta
\end{aligned}$$

or

$$\ddot{\theta}\ +\ \sin\theta\left(\frac{g}{R}\ -\ \Omega^2\ \cos\theta\right)\ =\ 0$$

Note that the support force provided by the hoop (necessary in the Newtonian method) is now replaced by the constraint $R =$ constant in the Lagrangian method. Furthermore, although the motion intrinsically takes place on the surface of a sphere of radius R, the azimuthal motion is completely determined by the equation $\varphi(t) = \varphi_0 + \Omega\, t$ and, thus, the motion of the bead takes place in one dimension.

Lastly, we note that this equation displays *bifurcation* behavior which is investigated in Chapter 8. For $\Omega^2 < g/R$, the equilibrium point $\theta = 0$ is stable while, for $\Omega^2 > g/R$, the equilibrium point $\theta = 0$ is now unstable and the new equilibrium point $\theta = \arccos(g/\Omega^2 R)$ is stable.

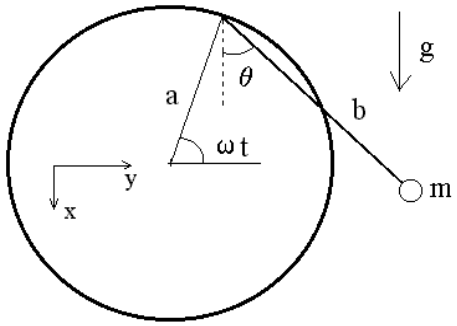


Figure 2.5: Generalized coordinates for the rotating-pendulum problem

2.3.3 Example III: Rotating Pendulum

As a third example, we consider a pendulum of mass m and length b attached to the edge of a disk of radius a rotating at angular velocity ω in a constant gravitational field with acceleration g. Placing the origin at the center of the disk, the coordinates of the pendulum mass are

$$\begin{aligned} x &= -a\,\sin\omega t \;+\; b\,\cos\theta \\ y &= a\,\cos\omega t \;+\; b\,\sin\theta \end{aligned}$$

so that the velocity components are

$$\begin{aligned} \dot{x} &= -a\omega\,\cos\omega t \;-\; b\,\dot{\theta}\,\sin\theta \\ \dot{y} &= -a\omega\,\sin\omega t \;+\; b\,\dot{\theta}\,\cos\theta \end{aligned}$$

and the squared velocity is

$$v^2 \;=\; a^2\omega^2 \;+\; b^2\dot{\theta}^2 \;+\; 2\,ab\,\omega\,\dot{\theta}\;\sin(\theta-\omega t).$$

Setting the zero potential energy at $x=0$, the gravitational potential energy is

$$U \;=\; -\,mg\,x \;=\; mga\,\sin\omega t \;-\; mgb\,\cos\theta.$$

The Lagrangian $L=K-U$ is, therefore, written as

$$\begin{aligned} L(\theta,\dot{\theta};t) \;=\; & \frac{m}{2}\left[\; a^2\omega^2 \;+\; b^2\dot{\theta}^2 \;+\; 2\,ab\,\omega\,\dot{\theta}\;\sin(\theta-\omega t)\;\right] \\ & -\; mga\,\sin\omega t \;+\; mgb\,\cos\theta, \end{aligned} \tag{2.13}$$

and the Euler-Lagrange equation for θ is

$$\begin{aligned} \frac{\partial L}{\partial\dot{\theta}} \;&=\; mb^2\,\dot{\theta} + m\,ab\,\omega\;\sin(\theta-\omega t) \;\;\rightarrow \\ \frac{d}{dt}\left(\frac{\partial L}{\partial\dot{\theta}}\right) \;&=\; mb^2\,\ddot{\theta} \;+\; m\,ab\,\omega\,(\dot{\theta}-\omega)\;\cos(\theta-\omega t) \end{aligned}$$

and

$$\frac{\partial L}{\partial \theta} \;=\; m\,ab\,\omega\,\dot{\theta}\,\cos(\theta-\omega t) \;-\; mg\,b\,\sin\theta$$

or

$$\ddot{\theta} \;+\; \frac{g}{b}\,\sin\theta \;-\; \frac{a}{b}\,\omega^2\,\cos(\theta-\omega t) \;=\; 0$$

We recover the standard equation of motion for the pendulum when a or ω vanish.

Note that the terms

$$\frac{m}{2}\,a^2\,\omega^2 \;-\; mga\,\sin\omega t$$

in the Lagrangian (2.13) play no role in determining the dynamics of the system. In fact, as can easily be shown, a Lagrangian L is always defined up to an exact time derivative, i.e., the Lagrangians L and $L' = L - df/dt$, where $f(\mathbf{q},t)$ is an arbitrary function, lead to the same Euler-Lagrange equations (see Section 2.4). In the present case,

$$f(t) \;=\; [(m/2)\,a^2\omega^2]\,t + (mga/\omega)\,\cos\omega t$$

and thus this term can be omitted from the Lagrangian (2.13) without changing the equations of motion.

2.3.4 Example IV: Compound Atwood Machine

As a fourth (and penultimate) example, we consider a compound Atwood machine composed three masses (labeled m_1, m_2, and m_3) attached by two massless ropes through two massless pulleys in a constant gravitational field with acceleration g.

The two generalized coordinates for this system (see Figure 2.6) are the distance x of mass m_1 from the top of the first pulley and the distance y of mass m_2 from the top of the second pulley; here, the lengths ℓ_a and ℓ_b are constants. The coordinates and velocities of the three masses m_1, m_2, and m_3 are

$$\begin{aligned} x_1 \;&=\; x \;\;\rightarrow\;\; v_1 \;=\; \dot{x}, \\ x_2 \;=\; \ell_a \;-\; x \;+\; y \;\;&\rightarrow\;\; v_2 \;=\; \dot{y} - \dot{x}, \\ x_3 \;=\; \ell_a \;-\; x \;+\; \ell_b \;-\; y \;\;&\rightarrow\;\; v_3 \;=\; -\dot{x} - \dot{y}, \end{aligned}$$

respectively, so that the total kinetic energy is

$$K \;=\; \frac{m_1}{2}\,\dot{x}^2 \;+\; \frac{m_2}{2}\,(\dot{y}-\dot{x})^2 \;+\; \frac{m_3}{2}\,(\dot{x}+\dot{y})^2\,.$$

Placing the zero potential energy at the top of the first pulley, the total gravitational potential energy, on the other hand, can be written as

$$U \;=\; -\,g\,x\;(m_1 - m_2 - m_3) \;-\; g\,y\;(m_2 - m_3)\,,$$

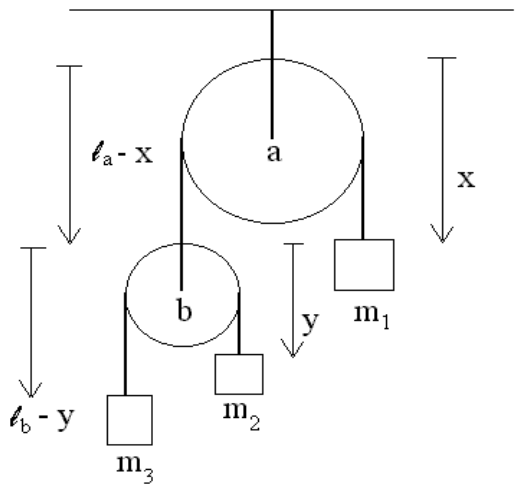


Figure 2.6: Generalized coordinates for the compound-Atwood problem

where constant terms were omitted. The Lagrangian $L = K - U$ is, therefore, written as

$$\begin{aligned} L(x, \dot{x},\ y, \dot{y}) \ &= \ \frac{m_1}{2}\,\dot{x}^2 \ + \ \frac{m_2}{2}\,(\dot{x} - \dot{y})^2 \ + \ \frac{m_3}{2}\,(\dot{x} + \dot{y})^2 \\ &\quad + \ g\,x\ (m_1 - m_2 - m_3) \ + \ g\,y\ (m_2 - m_3)\,. \end{aligned}$$

The Euler-Lagrange equation for x is

$$\begin{aligned} \frac{\partial L}{\partial \dot{x}} \ &= \ (m_1 + m_2 + m_3)\,\dot{x} + (m_3 - m_2)\,\dot{y} \ \ \rightarrow \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) \ &= \ (m_1 + m_2 + m_3)\,\ddot{x} + (m_3 - m_2)\,\ddot{y} \\ \frac{\partial L}{\partial x} \ &= \ g\ (m_1 - m_2 - m_3) \end{aligned}$$

while the Euler-Lagrange equation for y is

$$\begin{aligned} \frac{\partial L}{\partial \dot{y}} \ &= \ (m_3 - m_2)\,\dot{x} \ + \ (m_2 + m_3)\,\dot{y} \ \ \rightarrow \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{y}}\right) \ &= \ (m_3 - m_2)\,\ddot{x} \ + \ (m_2 + m_3)\,\ddot{y} \\ \frac{\partial L}{\partial y} \ &= \ g\ (m_2 - m_3)\,. \end{aligned}$$

We combine these two Euler-Lagrange equations

$$(m_1 + m_2 + m_3)\,\ddot{x} + (m_3 - m_2)\,\ddot{y} \ = \ g\ (m_1 - m_2 - m_3)\,, \tag{2.14}$$

$$(m_3 - m_2)\,\ddot{x} \ + \ (m_2 + m_3)\,\ddot{y} \ = \ g\ (m_2 - m_3)\,, \tag{2.15}$$

to describe the dynamical evolution of the Compound Atwood Machine. This set of equations can, in fact, be solved explicitly as

$$\ddot{x} \;=\; g\,\left(\frac{m_1\,m_+ \;-\; (m_+^2 - m_-^2)}{m_1\,m_+ \;+\; (m_+^2 - m_-^2)}\right) \quad \text{and} \quad \ddot{y} \;=\; g\,\left(\frac{2\,m_1\,m_-}{m_1\,m_+ \;+\; (m_+^2 - m_-^2)}\right),$$

where $m_\pm = m_2 \pm m_3$. Note also that it can be shown that the position z of the center of mass of the mechanical system (as measured from the top of the first pulley) satisfies the relation

$$Mg\,(z - z_0) \;=\; \frac{m_1}{2}\,\dot{x}^2 \;+\; \frac{m_2}{2}\,(\dot{y} - \dot{x})^2 \;+\; \frac{m_3}{2}\,(\dot{x} + \dot{y})^2\,, \tag{2.16}$$

where $M = (m_1 + m_2 + m_3)$ denotes the total mass of the system and we have assumed that the system starts from rest (with its center of mass located at z_0). This important relation tells us that, as the masses start to move, the center of mass must fall.

At this point, we introduce a convenient technique (henceforth known as *Freezing Degrees of Freedom*) for checking on the physical accuracy of any set of coupled Euler-Lagrange equations. Hence, for the Euler-Lagrange equation (2.14), we may freeze the degree of freedom associated with the y-coordinate (i.e., we set $\dot{y} = 0 = \ddot{y}$ or $m_- = 0$) to obtain

$$\ddot{x} \;=\; g\,\left(\frac{m_1 \;-\; m_+}{m_1 \;+\; m_+}\right),$$

in agreement with the analysis of a *simple* Atwood machine composed of a mass m_1 on one side and a mass $m_+ = m_2 + m_3$ on the other side. Likewise, for the Euler-Lagrange equation (2.15), we may freeze the degree of freedom associated with the x-coordinate (i.e., we set $\dot{x} = 0 = \ddot{x}$ or $m_1 m_+ = m_+^2 - m_-^2$) to obtain $\ddot{y} = g\,(m_-/m_+)$, again in agreement with the analysis of a *simple* Atwood machine.

2.3.5 Example V: Pendulum with Oscillating Fulcrum

As a fifth and final example, we consider the case of a pendulum of mass m and length ℓ attached to a massless block which is attached to a fixed wall by a massless spring of constant k; of course, we assume that the massless block moves without friction on a set of rails. Here, we use the two generalized coordinates x and θ shown in Figure 2.7 and write the Cartesian coordinates (y, z) of the pendulum mass as $y = x + \ell\,\sin\theta$ and $z = -\,\ell\,\cos\theta$, with its associated velocity components $v_y = \dot{x} + \ell\dot{\theta}\,\cos\theta$ and $v_z = \ell\dot{\theta}\,\sin\theta$. The kinetic energy of the pendulum is thus

$$K \;=\; \frac{m}{2}\left(v_y^2 \;+\; v_z^2\right) \;=\; \frac{m}{2}\left(\dot{x}^2 \;+\; \ell^2\dot{\theta}^2 \;+\; 2\,\ell\,\cos\theta\,\dot{x}\dot{\theta}\right).$$

The potential energy $U = U_k + U_g$ has two terms: one term $U_k = \frac{1}{2}\,kx^2$ associated with displacement of the spring away from its equilibrium position and one term $U_g = mgz$ associated with gravity. Hence, the Lagrangian for this system is

$$L(x, \theta, \dot{x}, \dot{\theta}) \;=\; \frac{m}{2}\left(\dot{x}^2 \;+\; \ell^2\dot{\theta}^2 \;+\; 2\,\ell\,\cos\theta\,\dot{x}\dot{\theta}\right) \;-\; \frac{k}{2}\,x^2 \;+\; mg\ell\,\cos\theta.$$

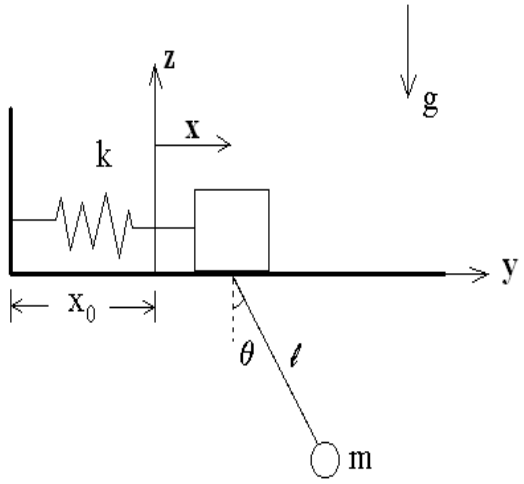


Figure 2.7: Generalized coordinates for the oscillating-pendulum problem

The Euler-Lagrange equation for x is

$$\begin{aligned}
\frac{\partial L}{\partial \dot{x}} &= m\left(\dot{x} + \ell\cos\theta\,\dot{\theta}\right) \;\rightarrow \\
\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) &= m\,\ddot{x} + m\ell\left(\ddot{\theta}\cos\theta - \dot{\theta}^2\sin\theta\right) \\
\frac{\partial L}{\partial x} &= -\,k\,x
\end{aligned}$$

while the Euler-Lagrange equation for θ is

$$\begin{aligned}
\frac{\partial L}{\partial \dot{\theta}} &= m\ell\left(\ell\,\dot{\theta} + \dot{x}\cos\theta\right) \;\rightarrow \\
\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) &= m\ell^2\,\ddot{\theta} + m\ell\left(\ddot{x}\cos\theta - \dot{x}\dot{\theta}\sin\theta\right) \\
\frac{\partial L}{\partial \theta} &= -\,m\ell\,\dot{x}\dot{\theta}\sin\theta - mg\ell\sin\theta
\end{aligned}$$

or

$$m\,\ddot{x} + k\,x = m\ell\left(\dot{\theta}^2\sin\theta - \ddot{\theta}\cos\theta\right), \tag{2.17}$$

$$\ddot{\theta} + (g/\ell)\sin\theta = -\,(\ddot{x}/\ell)\cos\theta. \tag{2.18}$$

Here, we recover the dynamical equation for a block-and-spring harmonic oscillator from Eq. (2.17) by freezing the degree of freedom associated with the θ-coordinate (i.e., by setting

$\dot{\theta} = 0 = \ddot{\theta}$) and the dynamical equation for the pendulum from Eq. (2.18) by freezing the degree of freedom associated with the x-coordinate (i.e., by setting $\dot{x} = 0 = \ddot{x}$). It is easy to see from this last example how powerful and yet simple the Lagrangian method is compared to the Newtonian method.

2.4 Symmetries and Conservation Laws

We are sometimes faced with a Lagrangian function which is either independent of time, independent of a linear spatial coordinate, or independent of an angular spatial coordinate. The Noether theorem (Amalie Emmy Noether, 1882-1935) states that *for each symmetry of the Lagrangian there corresponds a conservation law (and vice versa).* When the Lagrangian L is invariant under a time translation, a space translation, or a spatial rotation, the conservation law involves energy, linear momentum, or angular momentum, respectively.

We begin our discussion with a general expression for the variation δL of the Lagrangian $L(\mathbf{q}, \dot{\mathbf{q}}, t)$:

$$\delta L \;=\; \delta\mathbf{q}\cdot\left[\;\frac{\partial L}{\partial \mathbf{q}} \;-\; \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right)\;\right] \;+\; \frac{d}{dt}\left(\delta\mathbf{q}\cdot\frac{\partial L}{\partial \dot{\mathbf{q}}}\right),$$

obtained after re-arranging the term $\delta\dot{\mathbf{q}}\cdot\partial L/\partial\dot{\mathbf{q}}$. Next, we make use of the Euler-Lagrange equations for $\mathbf{q}$ (which enables us to drop the term $\delta\mathbf{q}\cdot[\cdots]$) and we find

$$\delta L \;=\; \frac{d}{dt}\left(\delta\mathbf{q}\cdot\frac{\partial L}{\partial \dot{\mathbf{q}}}\right).$$

Lastly, the variation δL can only be generated by a time translation δt, since

$$\begin{aligned} 0 \;=\; \delta\int L\,dt \;&=\; \int\left[\left(\delta L \;+\; \delta t\,\frac{\partial L}{\partial t}\right)dt \;+\; L\,d\delta t\;\right] \\ &=\; \int\left[\;\delta L \;-\; \delta t\left(\frac{dL}{dt} \;-\; \frac{\partial L}{\partial t}\right)\;\right]dt \end{aligned}$$

so that

$$\delta L \;=\; \delta t\left(\frac{dL}{dt} \;-\; \frac{\partial L}{\partial t}\right),$$

and, hence, we find

$$\delta t\left(\frac{dL}{dt} \;-\; \frac{\partial L}{\partial t}\right) \;=\; \frac{d}{dt}\left(\delta\mathbf{q}\cdot\frac{\partial L}{\partial \dot{\mathbf{q}}}\right), \tag{2.19}$$

which we, henceforth, refer to as the Noether equation for finite-dimensional mechanical systems [see Eq. (9.10) in Chapter 9 for the infinite-dimensional case].

2.4.1 Energy Conservation Law

We now apply the Noether equation (2.19) to investigate Noether's Theorem. First, we consider time translations, $t \to t + \delta t$ and $\delta \mathbf{q} = \dot{\mathbf{q}}\, \delta t$, so that the Noether equation (2.19) becomes

$$-\frac{\partial L}{\partial t} \;=\; \frac{d}{dt}\left(\dot{\mathbf{q}} \cdot \frac{\partial L}{\partial \dot{\mathbf{q}}} \;-\; L\right).$$

Noether's Theorem states that if the Lagrangian is invariant under time translations, i.e., $\partial L/\partial t = 0$, then energy is conserved, $dE/dt = 0$, where

$$E \;=\; \dot{\mathbf{q}} \cdot \frac{\partial L}{\partial \dot{\mathbf{q}}} \;-\; L$$

defines the energy invariant.

2.4.2 Momentum Conservation Law

Next, we consider invariance under spatial translations, $\mathbf{q} \to \mathbf{q} + \boldsymbol{\epsilon}$ (where $\delta \mathbf{q} = \boldsymbol{\epsilon}$ denotes a constant infinitesimal displacement), so that the Noether equation (2.19) yields the *linear* momentum conservation law

$$0 \;=\; \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{q}}}\right) \;=\; \frac{d\mathbf{P}}{dt},$$

where $\mathbf{P}$ denotes the total linear momentum of the mechanical system. On the other hand, when the Lagrangian is invariant under spatial rotations, $\mathbf{q} \to \mathbf{q} + (\delta \boldsymbol{\varphi} \times \mathbf{q})$ (where $\delta \boldsymbol{\varphi} = \delta\varphi\, \widehat{\varphi}$ denotes a constant infinitesimal rotation about an axis along the $\widehat{\varphi}$-direction), the Noether equation (2.19) yields the *angular* momentum conservation law

$$0 \;=\; \frac{d}{dt}\left(\mathbf{q} \times \frac{\partial L}{\partial \dot{\mathbf{q}}}\right) \;=\; \frac{d\mathbf{L}}{dt},$$

where $\mathbf{L} = \mathbf{q} \times \mathbf{P}$ denotes the total angular momentum of the mechanical system.

2.4.3 Invariance Properties

Lastly, an important invariance property of the Lagrangian is related to the fact that the Euler-Lagrange equations themselves are invariant under the transformation $L \to L + dF/dt$ on the Lagrangian itself, where $F(\mathbf{q}, t)$ is an arbitrary function. We call $L' = L + dF/dt$ the new Lagrangian and L the old Lagrangian. The Euler-Lagrange equations for the new Lagrangian are

$$\frac{d}{dt}\left(\frac{\partial L'}{\partial \dot{q}^i}\right) \;=\; \frac{\partial L'}{\partial q^i},$$

where

$$\frac{dF(\mathbf{q},t)}{dt} \;=\; \frac{\partial F}{\partial t} \;+\; \sum_j \dot{q}^j \,\frac{\partial F}{\partial q^j}.$$

Let us begin with

$$\frac{\partial L'}{\partial \dot{q}^i} \;=\; \frac{\partial}{\partial \dot{q}^i}\left(L \;+\; \frac{\partial F}{\partial t} \;+\; \sum_j \dot{q}^j \,\frac{\partial F}{\partial q^j}\right) \;=\; \frac{\partial L}{\partial \dot{q}^i} \;+\; \frac{\partial F}{\partial q^i},$$

so that

$$\frac{d}{dt}\left(\frac{\partial L'}{\partial \dot{q}^i}\right) \;=\; \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) \;+\; \frac{\partial^2 F}{\partial t \partial q^i} \;+\; \sum_k \dot{q}^k \,\frac{\partial^2 F}{\partial q^k \partial q^i}.$$

Next, we find

$$\frac{\partial L'}{\partial q^i} \;=\; \frac{\partial}{\partial q^i}\left(L \;+\; \frac{\partial F}{\partial t} \;+\; \sum_j \dot{q}^j \,\frac{\partial F}{\partial q^j}\right) \;=\; \frac{\partial L}{\partial q^i} \;+\; \frac{\partial^2 F}{\partial q^i \partial t} \;+\; \sum_j \dot{q}^j \,\frac{\partial^2 F}{\partial q^i \partial q^j}.$$

Using the symmetry properties

$$\dot{q}^j \,\frac{\partial^2 F}{\partial q^i \partial q^j} \;=\; \dot{q}^j \,\frac{\partial^2 F}{\partial q^j \partial q^i} \quad \text{and} \quad \frac{\partial^2 F}{\partial t \partial q^i} \;=\; \frac{\partial^2 F}{\partial q^i \partial t},$$

we easily verify

$$\frac{d}{dt}\left(\frac{\partial L'}{\partial \dot{q}^i}\right) \;-\; \frac{\partial L'}{\partial q^i} \;=\; \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) \;-\; \frac{\partial L}{\partial q^i} \;=\; 0,$$

and thus since L and $L' = L + dF/dt$ lead to the same Euler-Lagrange equations, they are said to be equivalent.

Using this invariance property, we note that the Galilean invariance of the Lagrangian $L(\mathbf{r},\mathbf{v})$ associated with velocity translations, $\mathbf{v} \to \mathbf{v} + \boldsymbol{\epsilon}$, yields the Lagrangian variation

$$\delta L \;=\; \boldsymbol{\epsilon} \cdot \;\left(\mathbf{v}\,\frac{\partial L}{\partial v^2}\right),$$

which, using the kinetic identity $\partial L/\partial v^2 = m/2$, can be written as an exact time derivative

$$\delta L \;=\; \frac{d}{dt}\left(\frac{\boldsymbol{\epsilon}}{2} \;\cdot\; m\mathbf{r}\right) \;=\; \frac{df}{dt},$$

and, thus, can be eliminated from the system Lagrangian.

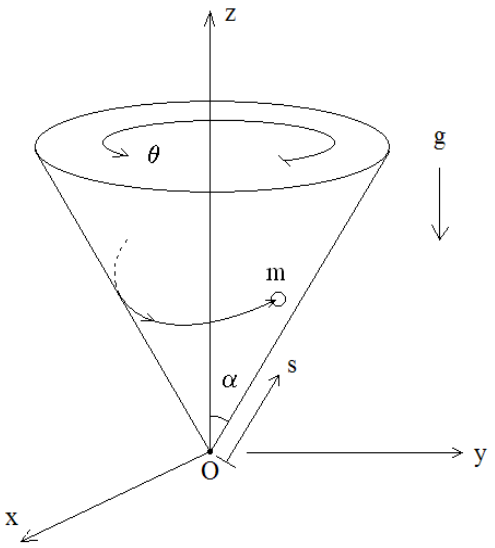


Figure 2.8: Motion on the surface of a cone

2.4.4 Lagrangian Mechanics with Symmetries

As an example of Lagrangian mechanics with symmetries, we return to the motion of a particle of mass m constrained to move on the surface of a cone of apex angle α (such that $r = z \tan\alpha$) in the presence of a gravitational field (see Figure 2.8 and Sec. 2.2.2). The Lagrangian for this constrained mechanical system is expressed in terms of the generalized coordinates (s, θ), where s denotes the distance from the cone's apex (labeled O in Figure 2.8) and θ is the standard polar angle in the (x, y)-plane. Hence, by combining the kinetic energy $K = \frac{1}{2} m(\dot{s}^2 + s^2 \dot{\theta}^2 \ \sin^2\alpha)$ with the potential energy $U = mgz = mg\, s \ \cos\alpha$, we construct the Lagrangian

$$L(s, \theta; \dot{s}, \dot{\theta}) \ = \ \frac{1}{2} m \left(\dot{s}^2 + s^2 \, \dot{\theta}^2 \ \sin^2\alpha \right) \ - \ mg\, s \ \cos\alpha. \tag{2.20}$$

Since the Lagrangian is independent of the polar angle θ, the canonical angular momentum

$$p_\theta \ = \ \frac{\partial L}{\partial \dot{\theta}} \ = \ ms^2 \, \dot{\theta} \ \sin^2\alpha \tag{2.21}$$

is a constant of the motion (as predicted by Noether's Theorem). The Euler-Lagrange equation for s, on the other hand, is expressed as

$$\ddot{s} \ + \ g \ \cos\alpha \ = \ s \, \dot{\theta}^2 \ \sin^2\alpha \ = \ \frac{\ell^2}{m^2 \, s^3 \sin^2\alpha}, \tag{2.22}$$

where $g \cos\alpha$ denotes the component of the gravitational acceleration parallel to the surface of the cone and $\ell = p_\theta$ denotes the constant value of the angular momentum. The right

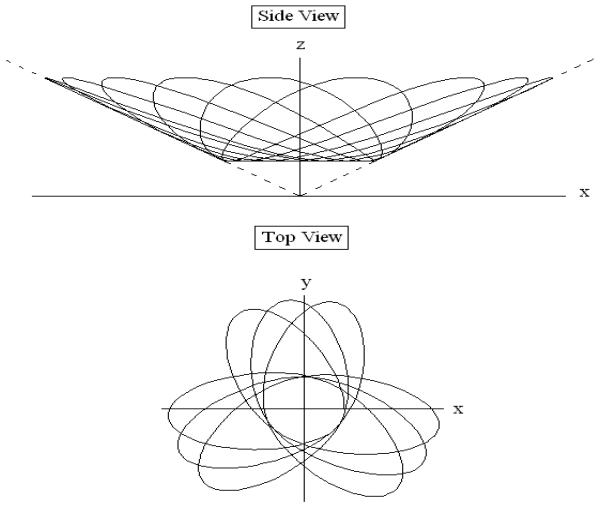


Figure 2.9: Particle orbits on the surface of a cone

side of Eq. (2.22) involves s only after using $\dot{\theta} = \ell/(m\, s^2 \sin^2\alpha)$, which follows from the conservation of angular momentum.

Figure 2.9 shows the results of the numerical integration of the Euler-Lagrange equations (2.21)-(2.22) for $\theta(t)$ and $s(t)$. The top figure in Figure 2.9 shows a projection of the path of the particle on the (x, z)-plane (side view), which clearly shows that the motion is periodic as the s-coordinate oscillates between two finite values of s. The bottom figure in Figure 2.9 shows a projection of the path of the particle on the (x, y)-plane (top view), which shows the slow precession motion in the θ-coordinate. In the next Chapter, we will show that the doubly-periodic motion of the particle is a result of the conservation law of angular momentum and energy (since the Lagrangian system is also independent of time).

2.4.5 Routh's Procedure for Eliminating Ignorable Coordinates

Edward John Routh (1831-1907) introduced a simple procedure for eliminating ignorable degrees of freedom while introducing their corresponding conserved momenta. Consider, for example, two-dimensional motion on the (x, y)-plane represented by the Lagrangian $L(r;\, \dot{r}, \dot{\theta})$, where r and θ are the polar coordinates. Since the Lagrangian under consideration is independent of the angle θ, the canonical momentum $p_\theta = \partial L/\partial\dot{\theta}$ is conserved. Routh's procedure for deriving a *reduced* Lagrangian involves the construction of the *Routh-Lagrange* function $R(r, \dot{r};\, p_\theta)$ defined as

$$R(r, \dot{r};\, p_\theta) \;=\; L(r;\, \dot{r}, \dot{\theta}) \;-\; p_\theta\, \dot{\theta}, \tag{2.23}$$

where $\dot{\theta}$ is expressed as a function of r and p_θ.

For example, for the constrained motion of a particle on the surface of a cone in the presence of gravity, the Lagrangian (2.20) can be reduced to the *Routh-Lagrange* function

$$R(s, \dot{s}; p_\theta) \;=\; \frac{1}{2}\, m\dot{s}^2 \;-\; \left(mg\, s\, \cos\alpha \;+\; \frac{\ell^2}{2m^2 s^2\, \sin^2\alpha} \right) \;=\; \frac{1}{2}\, m\dot{s}^2 \;-\; V(s), \qquad (2.24)$$

and the equation of motion (2.22) can be expressed in Euler-Lagrange form

$$\frac{d}{ds}\left(\frac{\partial R}{\partial \dot{s}} \right) \;=\; \frac{\partial R}{\partial s} \quad \rightarrow \quad m\, \ddot{s} \;=\; -\, V'(s),$$

in terms of the effective potential

$$V(s) \;=\; mg\, s\, \cos\alpha \;+\; \frac{\ell^2}{2m^2 s^2\, \sin^2\alpha}.$$

2.5 Lagrangian Mechanics in the Center-of-Mass Frame

An important frame of reference associated with the dynamical description of the motion of several interacting particles is provided by the center-of-mass (CM) frame. The following discussion focuses on the Lagrangian for an isolated two-particle system expressed as

$$L \;=\; \frac{m_1}{2}\, |\dot{\mathbf{r}}_1|^2 \;+\; \frac{m_2}{2}\, |\dot{\mathbf{r}}_2|^2 \;-\; U(\mathbf{r}_1 - \mathbf{r}_2),$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ represent the positions of the particles of mass m_1 and m_2, respectively, and $U(\mathbf{r}_1, \mathbf{r}_2) = U(\mathbf{r}_1 - \mathbf{r}_2)$ is the potential energy for an isolated two-particle system (see Figure below).

Let us now define the position $\mathbf{R}$ of the center of mass

$$\mathbf{R} \;=\; \frac{m_1\, \mathbf{r}_1 + m_2\, \mathbf{r}_2}{m_1 + m_2},$$

and define the inter-particle vector $\mathbf{r} \;=\; \mathbf{r}_1 - \mathbf{r}_2$, so that the particle positions can be expressed as

$$\mathbf{r}_1 \;=\; \mathbf{R} \;+\; \frac{m_2}{M}\, \mathbf{r} \quad \text{and} \quad \mathbf{r}_2 \;=\; \mathbf{R} \;-\; \frac{m_1}{M}\, \mathbf{r},$$

where $M = m_1 + m_2$ is the total mass of the two-particle system (see Figure 2.10). The Lagrangian of the isolated two-particle system thus becomes

$$L \;=\; \frac{M}{2}\, |\dot{\mathbf{R}}|^2 \;+\; \frac{\mu}{2}\, |\dot{\mathbf{r}}|^2 \;-\; U(\mathbf{r}),$$

where

$$\mu \;=\; \frac{m_1\, m_2}{m_1 + m_2} \;=\; \left(\frac{1}{m_1} \;+\; \frac{1}{m_2} \right)^{-1}$$

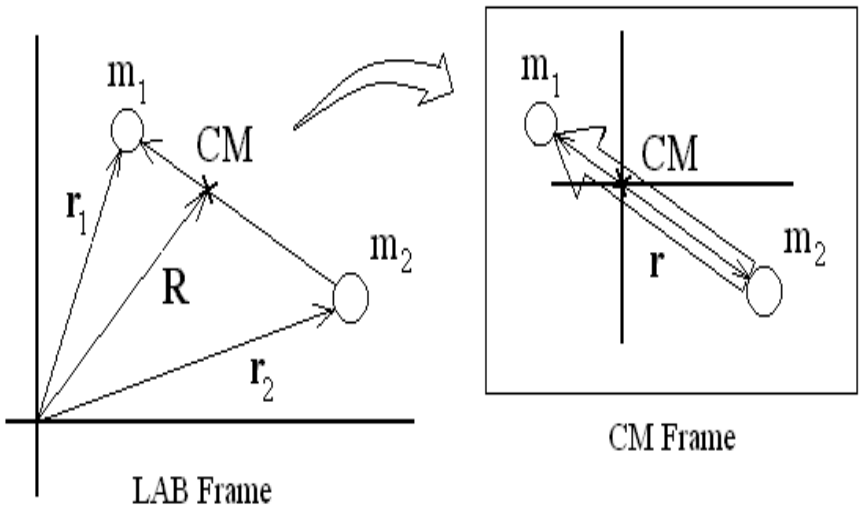


Figure 2.10: Center-of-Mass frame

denotes the *reduced* mass of the two-particle system. We note that the angular momentum of the two-particle system is expressed as

$$\mathbf{L} \;=\; \sum_a \mathbf{r}_a \times \mathbf{p}_a \;=\; \mathbf{R}\times\mathbf{P} \;+\; \mathbf{r}\times\mathbf{p}, \tag{2.25}$$

where the canonical momentum of the center-of-mass $\mathbf{P}$ and the canonical momentum $\mathbf{p}$ of the two-particle system in the CM frame are defined, respectively, as

$$\mathbf{P} \;=\; \frac{\partial L}{\partial \dot{\mathbf{R}}} \;=\; M\,\dot{\mathbf{R}} \quad \text{and} \quad \mathbf{p} \;=\; \frac{\partial L}{\partial \dot{\mathbf{r}}} \;=\; \mu\dot{\mathbf{r}}.$$

For an isolated system, the canonical momentum $\mathbf{P}$ of the center-of-mass is a constant of the motion. The CM reference frame is defined by the condition $\mathbf{R} = 0$, i.e., we move the origin of our coordinate system to the CM position.

In the CM frame, the Lagrangian for an isolated two-particle system in the CM reference frame

$$L(\mathbf{r},\,\dot{\mathbf{r}}) \;=\; \frac{\mu}{2}\,|\dot{\mathbf{r}}|^2 \;-\; U(\mathbf{r}), \tag{2.26}$$

describes the motion of a *fictitious* particle of mass μ at position $\mathbf{r}$, where the positions of the two *real* particles of masses m_1 and m_2 are

$$\mathbf{r}_1 \;=\; \frac{m_2}{M}\,\mathbf{r} \quad \text{and} \quad \mathbf{r}_2 \;=\; -\,\frac{m_1}{M}\,\mathbf{r}. \tag{2.27}$$

Hence, once the Euler-Lagrange equation for $\mathbf{r}$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{r}}}\right) \;=\; \frac{\partial L}{\partial \mathbf{r}} \quad\rightarrow\quad \mu\,\ddot{\mathbf{r}} \;=\; -\,\nabla U(\mathbf{r})$$

is solved, the motion of the two particles is determined through Eqs. (2.27). The angular momentum $\mathbf{L} = \mu\,\mathbf{r} \times \dot{\mathbf{r}}$ in the CM frame satisfies the evolution equation

$$\frac{d\mathbf{L}}{dt} \;=\; \mathbf{r} \times \mu\ddot{\mathbf{r}} \;=\; -\,\mathbf{r} \times \nabla U(\mathbf{r}). \tag{2.28}$$

Here, using spherical coordinates (r, θ, φ), we find

$$\frac{d\mathbf{L}}{dt} \;=\; [\,\cot\theta\,(\cos\varphi\,\hat{\mathsf{x}} + \sin\varphi\,\hat{\mathsf{y}}) \;-\; \hat{\mathsf{z}}\,]\;\frac{\partial U}{\partial \varphi} \;+\; (\sin\varphi\,\hat{\mathsf{x}} - \cos\varphi\,\hat{\mathsf{y}})\;\frac{\partial U}{\partial \theta}.$$

Hence, if motion is originally taking place on the (x, y)-plane (i.e., at $\theta = \pi/2$) and the potential $U(r, \varphi)$ is independent of the polar angle θ, then the angular momentum vector is $\mathbf{L} = \ell\,\hat{\mathsf{z}}$ and its magnitude ℓ satisfies the evolution equation

$$\frac{d\ell}{dt} \;=\; -\,\frac{\partial U}{\partial \varphi}.$$

Hence, for motion in a potential $U(r)$ that depends only on the radial position r, the angular momentum $\mathbf{L} = \ell\;\hat{\mathsf{z}}$ represents an additional constant of motion. Motion in such potentials is refered to as motion in a *central-force* potential and will be studied in Chap. 4.

2.6 Problems

Problem 1

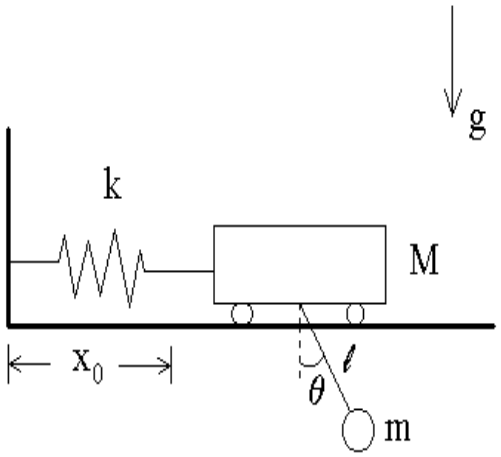
A particle of mass m is constrained to slide down a curve $y = V(x)$ under the action of gravity without friction. Show that the Euler-Lagrange equation for this system yields the equation

$$\ddot{x} \;=\; -\,V'\,\left(g \;+\; \ddot{V}\right),$$

where $\dot{V} = \dot{x}\,V'$ and $\ddot{V} = (\dot{V})^{\cdot} = \ddot{x}\,V' + \dot{x}^2\,V''$.

Problem 2

A cart of mass M is placed on rails and attached to a wall with the help of a massless spring with constant k (as shown in the Figure below); the spring is in its equilibrium state when the cart is at a distance x_0 from the wall. A pendulum of mass m and length ℓ is attached to the cart (as shown).

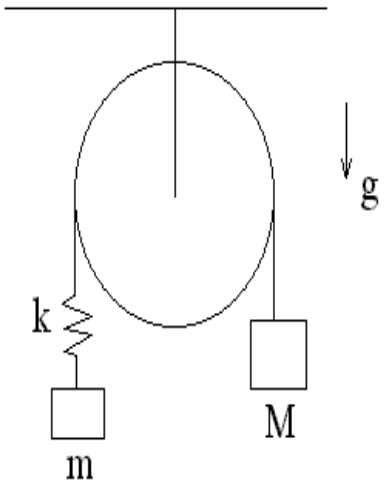


(a) Write the Lagrangian $L(x, \dot{x}, \theta, \dot{\theta})$ for the cart-pendulum system, where x denotes the position of the cart (as measured from a suitable origin) and θ denotes the angular position of the pendulum.

(b) From your Lagrangian, write the Euler-Lagrange equations for the generalized coordinates x and θ.

Problem 3

An Atwood machine is composed of two masses m and M attached by means of a massless rope into which a massless spring (with constant k) is inserted (as shown in the Figure below). When the spring is in a relaxed state, the spring-rope length is ℓ.



(a) Find suitable generalized coordinates to describe the motion of the two masses (allowing for elongation or compression of the spring).

(b) Using these generalized coordinates, construct the Lagrangian and derive the appropriate Euler-Lagrange equations.

Chapter 3

Hamiltonian Mechanics

3.1 Canonical Hamilton's Equations

In the previous Chapter, the Lagrangian method was introduced as a powerful alternative to the Newtonian method for deriving equations of motion for complex mechanical systems. In the present Chapter, a complementary approach to the Lagrangian method, known as the Hamiltonian method, is presented. Although much of the Hamiltonian method is outside the scope of this course (e.g., the Hamiltonian formulation of Quantum Mechanics), a simplified version (the Energy method) is presented as a powerful method for *solving* the Euler-Lagrange equations.

The k second-order Euler-Lagrange equations on configuration space $\mathbf{q} = (q^1, ..., q^k)$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^j}\right) \;=\; \frac{\partial L}{\partial q^j}, \tag{3.1}$$

can be written as $2k$ first-order differential equations, known as Hamilton's equations (William Rowan Hamilton, 1805-1865), on a $2k$-dimensional *phase* space $\mathsf{z} = (\mathbf{q};\ \mathbf{p})$ as

$$\frac{d\mathbf{q}}{dt} \;=\; \frac{\partial H}{\partial \mathbf{p}} \quad \text{and} \quad \frac{d\mathbf{p}}{dt} \;=\; -\,\frac{\partial H}{\partial \mathbf{q}}, \tag{3.2}$$

where

$$p_j(\mathbf{q}, \dot{\mathbf{q}}; t) \;=\; \frac{\partial L}{\partial \dot{q}^j}(\mathbf{q}, \dot{\mathbf{q}}; t) \tag{3.3}$$

defines the j^{th}-component of the canonical momentum. Here, the Hamiltonian function $H(\mathbf{q}, \mathbf{p}; t)$ is defined from the Lagrangian function $L(\mathbf{q}, \dot{\mathbf{q}}; t)$ by the Legendre transformation (Adrien-Marie Legendre, 1752-1833)

$$H(\mathbf{q}, \mathbf{p}; t) \;=\; \mathbf{p} \cdot \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t) \;-\; L[\mathbf{q}, \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t), t]. \tag{3.4}$$

We note that the converse of the Legendre transformation (3.4),

$$L(\mathbf{z}, \dot{\mathbf{z}}; t) \;=\; \mathbf{p} \cdot \dot{\mathbf{q}} - H(\mathbf{z}; t),$$

can be used in the variational principle

$$\delta \int\, L(\mathbf{q}, \mathbf{p}; t)\, dt \;=\; \int \left[\, \delta \mathbf{p} \cdot \left(\dot{\mathbf{q}} \;-\; \frac{\partial H}{\partial \mathbf{p}} \right) \;-\; \delta \mathbf{q} \cdot \left(\dot{\mathbf{p}} \;+\; \frac{\partial H}{\partial \mathbf{q}} \right) \,\right] \, dt \;=\; 0$$

to obtain Hamilton's equations (3.2) in the $2k$-dimensional *phase space* with coordinates $\mathbf{z} = (\mathbf{q}, \mathbf{p})$.

3.2 Legendre Transformation

Before proceeding with the Hamiltonian formulation of particle dynamics, we investigate the conditions under which the Legendre transformation (3.4) is possible. Once again, the Legendre transformation allows the transformation from a Lagrangian description of a dynamical system in terms of a Lagrangian function $L(\mathbf{r}, \dot{\mathbf{r}}, t)$ to a Hamiltonian description of the same dynamical system in terms of a Hamiltonian function $H(\mathbf{r}, \mathbf{p}, t)$, where the canonical momentum $\mathbf{p}$ is defined as $p_i = \partial L / \partial \dot{x}^i$.

It turns out that the condition under which the Legendre transformation can be used is associated with the condition under which the inversion of the relation $\mathbf{p}(\mathbf{r}, \dot{\mathbf{r}}, t) \rightarrow \dot{\mathbf{r}}(\mathbf{r}, \mathbf{p}, t)$ is possible. To simplify our discussion, we focus on motion in two dimensions (labeled x and y). The general expression of the kinetic energy term of a Lagrangian with two degrees of freedom $L(x, \dot{x},\; y, \dot{y}) = K(x, \dot{x},\; y, \dot{y}) - U(x, y)$ is

$$K(x, \dot{x},\; y, \dot{y}) \;=\; \frac{\alpha}{2}\, \dot{x}^2 \;+\; \beta\, \dot{x}\, \dot{y} \;+\; \frac{\gamma}{2}\, \dot{y}^2 \;=\; \frac{1}{2}\, \dot{\mathbf{r}}^T \;\cdot\; \mathsf{M} \;\cdot\; \dot{\mathbf{r}},$$

where $\dot{\mathbf{r}}^T = (\dot{x}, \dot{y})$ and the *mass* matrix M is

$$\mathsf{M} \;=\; \begin{pmatrix} \alpha & \beta \\ \\ \beta & \gamma \end{pmatrix}.$$

Here, the coefficients α, β, and γ may be function of x and/or y. The canonical momentum vector (3.3) is thus defined as

$$\mathbf{p} \;=\; \frac{\partial L}{\partial \dot{\mathbf{r}}} \;=\; \mathsf{M} \;\cdot\; \dot{\mathbf{r}} \;\;\rightarrow\;\; \begin{pmatrix} p_x \\ \\ p_y \end{pmatrix} \;=\; \begin{pmatrix} \alpha & \beta \\ \\ \beta & \gamma \end{pmatrix} \cdot \begin{pmatrix} \dot{x} \\ \\ \dot{y} \end{pmatrix}$$

or

$$\left.\begin{aligned} p_x \;&=\; \alpha\, \dot{x} \;+\; \beta\, \dot{y} \\ \\ p_y \;&=\; \beta\, \dot{x} \;+\; \gamma\, \dot{y} \end{aligned}\right\}. \tag{3.5}$$

The Lagrangian is said to be *regular* if the matrix M is invertible, i.e., if its determinant

$$\Delta \;=\; \alpha\,\gamma \;-\; \beta^2 \;\neq\; 0.$$

In the case of a regular Lagrangian, we readily invert (3.5) to obtain

$$\dot{\mathbf{r}}(\mathbf{r},\mathbf{p},t) \;=\; \mathsf{M}^{-1}\;\cdot\;\mathbf{p} \;\;\rightarrow\;\; \begin{pmatrix} \dot{x} \\ \\ \dot{y} \end{pmatrix} \;=\; \frac{1}{\Delta}\begin{pmatrix} \gamma & -\,\beta \\ \\ -\,\beta & \alpha \end{pmatrix}\cdot\begin{pmatrix} p_x \\ \\ p_y \end{pmatrix}$$

or

$$\left.\begin{aligned} \dot{x} \;&=\; (\gamma\,p_x \;-\; \beta\,p_y)/\Delta \\ \dot{y} \;&=\; (\alpha\,p_y \;-\; \beta\,p_x)/\Delta \end{aligned}\right\}, \tag{3.6}$$

and the kinetic energy term becomes

$$K(x,p_x,\;y,p_y) \;=\; \frac{1}{2}\,\mathbf{p}^T\;\cdot\;\mathsf{M}^{-1}\;\cdot\;\mathbf{p}.$$

Lastly, under the Legendre transformation, we find

$$\begin{aligned} H \;&=\; \mathbf{p}^T\;\cdot\;\Big(\mathsf{M}^{-1}\;\cdot\;\mathbf{p}\Big) \;-\; \Big(\frac{1}{2}\,\mathbf{p}^T\;\cdot\;\mathsf{M}^{-1}\;\cdot\;\mathbf{p} \;-\; U\Big) \\ &=\; \frac{1}{2}\,\mathbf{p}^T\;\cdot\;\mathsf{M}^{-1}\;\cdot\;\mathbf{p} \;+\; U. \end{aligned}$$

Hence, we clearly see that the Legendre transformation is applicable only if the mass matrix M is invertible.

Lastly, we note that the Legendre transformation is also used in Thermodynamics. Indeed, we begin with the First Law of Thermodynamics $dU(S,V) = T\,dS - P\,dV$ expressed in terms of the internal energy function $U(S,V)$, where entropy S and volume V are the independent variables while temperature $T(S,V) = \partial U/\partial S$ and pressure $P(S,V) = -\,\partial U/\partial V$ are dependent variables. It is possible, however, to choose other independent variables by defining new thermodynamic functions as shown in the Table below.

	Pressure P	Volume V	Temperature T	Entropy S
Pressure P	N/A	•	$G \;=\; H - T\,S$	$H \;=\; U + P\,V$
Volume V	•	N/A	$F \;=\; U - T\,S$	U
Temperature T	$G \;=\; H - T\,S$	$F \;=\; U - T\,S$	N/A	•
Entropy S	$H \;=\; U + P\,V$	U	•	N/A

For example, if we choose volume V and temperature T as independent variables, we introduce the Legendre transformation from the internal energy $U(S,V)$ to the Helmholtz free energy $F(V,T) = U - T\,S$, such that the First Law of Thermodynamics now becomes

$$dF(V,T) \;=\; dU \;-\; T\;dS \;-\; S\;dT \;=\; -\;P\;dV \;-\; S\;dT,$$

where pressure $P(V,T) = -\partial F/\partial V$ and entropy $S(V,T) = -\partial F/\partial T$ are dependent variables. Likewise, enthalpy $H(P,S) = U + PV$ and Gibbs free energy $G(T,P) = H - TS$ are introduced by Legendre transformations whenever one chooses (P,S) and (T,P), respectively, as independent variables.

3.3 Hamiltonian Optics and Wave-Particle Duality*

Historically, the Hamiltonian method was first introduced as a formulation of the dynamics of light rays. Consider the following *phase* integral

$$\Theta[\mathbf{z}] \;=\; \int_{t_1}^{t_2} \;[\; \mathbf{k}\cdot\dot{\mathbf{x}} \;-\; \omega(\mathbf{x},\mathbf{k};t)\;]\,dt, \tag{3.7}$$

where $\Theta[\mathbf{z}]$ is a functional of the path $\mathbf{z}(t) = (\mathbf{x}(t), \mathbf{k}(t))$ in *ray phase space*, expressed in terms of the instantaneous position $\mathbf{x}(t)$ of a light ray and its associated instantaneous wave vector $\mathbf{k}(t)$; here, the dispersion relation $\omega(\mathbf{x},\mathbf{k};t)$ is obtained as a root of the dispersion equation $\det \mathsf{D}(\mathbf{x},t;\,\mathbf{k},\omega) = 0$, and a dot denotes a total time derivative: $\dot{\mathbf{x}} = d\mathbf{x}/dt$.

Assuming that the phase integral $\Theta[\mathbf{z}]$ acquires a *minimal* value for a *physical* ray orbit $\mathbf{z}(t)$, henceforth called the Principle of *Phase Stationarity* $\delta\Theta = 0$, we can show that Euler's First Equation lead to *Hamilton's* ray equations

$$\frac{d\mathbf{x}}{dt} \;=\; \frac{\partial\omega}{\partial\mathbf{k}} \quad\text{and}\quad \frac{d\mathbf{k}}{dt} \;=\; -\,\nabla\omega. \tag{3.8}$$

The first ray equation states that a ray travels at the *group* velocity while the second ray equation states that the wave vector $\mathbf{k}$ is refracted as the ray propagates in a non-uniform medium (see Chapter 1). Hence, the frequency function $\omega(\mathbf{x},\mathbf{k};t)$ is the Hamiltonian of ray dynamics in a nonuniform medium.

It was Prince Louis Victor Pierre Raymond de Broglie (1892-1987) who noted (as a graduate student well versed in Classical Mechanics) the similarities between Hamilton's equations (3.2) and (3.8), on the one hand, and the Maupertuis-Jacobi (2.1) and Euler-Lagrange (2.8) Principles of Least Action and Fermat's Principle of Least Time (1.1) and Principle of Phase Stationarity (3.7), on the other hand. By using the *quantum of action* $\hbar = h/2\pi$ defined in terms of Planck's constant h and Planck's energy hypothesis $E = \hbar\omega$, de Broglie suggested that a particle's momentum $\mathbf{p}$ be related to its wavevector $\mathbf{k}$ according to de Broglie's formula $\mathbf{p} = \hbar\,\mathbf{k}$ and introduced the wave-particle synthesis based on the identity $S[\mathbf{z}] = \hbar\,\Theta[\mathbf{z}]$ involving the action integral $S[\mathbf{z}]$ and the phase integral $\Theta[\mathbf{z}]$:

	Particle	Wave
phase space	$\mathbf{z} = (\mathbf{q},\mathbf{p})$	$\mathbf{z} = (\mathbf{x},\mathbf{k})$
Hamiltonian	$H(\mathbf{z})$	$\omega(\mathbf{z})$
Variational Principle I	Maupertuis − Jacobi	Fermat
Variational Principle II	Euler − Lagrange	Phase − Stationarity

The final synthesis came with Richard Philips Feynman (1918-1988) who provided a derivation of Schroedinger's equation by associating the probability that a particle follow a particular path with the expression

$$\exp\left(\frac{i}{\hbar}\, S[\mathbf{z}]\right)$$

where $S[\mathbf{z}]$ denotes the action integral for the path (see Appendix A).

3.4 Particle Motion in an Electromagnetic Field*

Single-particle motion in an electromagnetic field represents the paradigm to illustrate the connection between Lagrangian and Hamiltonian mechanics.

3.4.1 Euler-Lagrange Equations

The equations of motion for a charged particle of mass m and charge e moving in an electromagnetic field represented by the electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$ are

$$\frac{d\mathbf{x}}{dt} \;=\; \mathbf{v} \tag{3.9}$$

$$\frac{d\mathbf{v}}{dt} \;=\; \frac{e}{m}\left(\mathbf{E} \,+\, \frac{d\mathbf{x}}{dt}\times\frac{\mathbf{B}}{c}\right), \tag{3.10}$$

where $\mathbf{x}$ denotes the position of the particle and $\mathbf{v}$ its velocity.

By treating the coordinates $(\mathbf{x}, \mathbf{v})$ as generalized coordinates (i.e., $\delta\mathbf{v}$ is independent of $\delta\mathbf{x}$), we now show that the equations of motion (3.9) and (3.10) can be obtained as Euler-Lagrange equations from the Lagrangian

$$L(\mathbf{x}, \dot{\mathbf{x}},\ \mathbf{v}, \dot{\mathbf{v}};\, t) \;=\; \left(m\,\mathbf{v} \,+\, \frac{e}{c}\,\mathbf{A}(\mathbf{x}, t)\right)\cdot\dot{\mathbf{x}} \;-\; \left(e\,\Phi(\mathbf{x}, t) \,+\, \frac{m}{2}\,|\mathbf{v}|^2\right), \tag{3.11}$$

where Φ and $\mathbf{A}$ are the electromagnetic potentials in terms of which electric and magnetic fields are defined

$$\mathbf{E} \;=\; -\,\nabla\Phi \;-\; \frac{1}{c}\,\frac{\partial\mathbf{A}}{\partial t} \;\text{ and }\; \mathbf{B} \;=\; \nabla\times\mathbf{A}. \tag{3.12}$$

Note that these expressions for $\mathbf{E}$ and $\mathbf{B}$ satisfy Faraday's law $\nabla\times\mathbf{E} \;=\; -\,c^{-1}\,\partial_t\mathbf{B}$ and Gauss' law $\nabla\cdot\mathbf{B} = 0$.

First, we look at the Euler-Lagrange equation for $\mathbf{x}$:

$$\frac{\partial L}{\partial\dot{\mathbf{x}}} \;=\; m\,\mathbf{v} \,+\, \frac{e}{c}\,\mathbf{A} \;\;\rightarrow\;\; \frac{d}{dt}\left(\frac{\partial L}{\partial\dot{\mathbf{x}}}\right) \;=\; m\,\dot{\mathbf{v}} \,+\, \frac{e}{c}\left(\frac{\partial\mathbf{A}}{\partial t} \,+\, \dot{\mathbf{x}}\cdot\nabla\mathbf{A}\right)$$

$$\frac{\partial L}{\partial\mathbf{x}} \;=\; \frac{e}{c}\,\nabla\mathbf{A}\cdot\dot{\mathbf{x}} \;-\; e\,\nabla\Phi,$$

which yields Eq. (3.10), since

$$m\,\dot{\mathbf{v}} \;=\; -\,e\left(\nabla\Phi \;+\; \frac{1}{c}\,\frac{\partial\mathbf{A}}{\partial t}\right) \;+\; \frac{e}{c}\,\dot{\mathbf{x}}\times\nabla\times\mathbf{A} \;=\; e\,\mathbf{E} \;+\; \frac{e}{c}\,\dot{\mathbf{x}}\times\mathbf{B}, \tag{3.13}$$

where the definitions (3.12) were used.

Next, we look at the Euler-Lagrange equation for $\mathbf{v}$:

$$\frac{\partial L}{\partial\dot{\mathbf{v}}} \;=\; 0 \;\;\rightarrow\;\; \frac{d}{dt}\left(\frac{\partial L}{\partial\dot{\mathbf{x}}}\right) \;=\; 0 \;=\; \frac{\partial L}{\partial\mathbf{v}} \;=\; m\,\dot{\mathbf{x}} \;-\; m\,\mathbf{v},$$

which yields Eq. (3.9). Because $\partial L/\partial\dot{\mathbf{v}} = 0$, we note that we could use Eq. (3.9) as a constraint which could be imposed *a priori* on the Lagrangian (3.11) to give

$$L(\mathbf{x},\dot{\mathbf{x}};\,t) \;=\; \frac{m}{2}\;|\dot{\mathbf{x}}|^2 \;+\; \frac{e}{c}\,\mathbf{A}(\mathbf{x},t)\cdot\dot{\mathbf{x}} \;-\; e\,\Phi(\mathbf{x},t). \tag{3.14}$$

The Euler-Lagrange equation in this case is identical to Eq. (3.13) with $\dot{\mathbf{v}} = \ddot{\mathbf{x}}$.

3.4.2 Energy Conservation Law

We now show that the second Euler equation (i.e., the energy conservation law), expressed as

$$\frac{d}{dt}\left(L \;-\; \dot{\mathbf{x}}\cdot\frac{\partial L}{\partial\dot{\mathbf{x}}} \;-\; \dot{\mathbf{v}}\cdot\frac{\partial L}{\partial\dot{\mathbf{v}}}\right) \;=\; \frac{\partial L}{\partial t},$$

is satisfied exactly by the Lagrangian (3.11) and the equations of motion (3.9) and (3.10).

First, from the Lagrangian (3.11), we find

$$\begin{aligned}
\frac{\partial L}{\partial t} \;&=\; \frac{e}{c}\,\frac{\partial\mathbf{A}}{\partial t}\cdot\dot{\mathbf{x}} \;-\; e\,\frac{\partial\Phi}{\partial t} \\
L \;-\; \dot{\mathbf{x}}\cdot\frac{\partial L}{\partial\dot{\mathbf{x}}} \;-\; \dot{\mathbf{v}}\cdot\frac{\partial L}{\partial\dot{\mathbf{v}}} \;&=\; L \;-\; \left(m\,\mathbf{v} \;+\; \frac{e}{c}\,\mathbf{A}\right)\cdot\dot{\mathbf{x}} \\
&=\; -\left(\frac{m}{2}\;|\mathbf{v}|^2 \;+\; e\,\Phi\right).
\end{aligned}$$

Next, we find

$$\frac{d}{dt}\left(L \;-\; \dot{\mathbf{x}}\cdot\frac{\partial L}{\partial\dot{\mathbf{x}}} \;-\; \dot{\mathbf{v}}\cdot\frac{\partial L}{\partial\dot{\mathbf{v}}}\right) \;=\; -\,m\mathbf{v}\cdot\dot{\mathbf{v}} \;-\; e\left(\frac{\partial\Phi}{\partial t} \;+\; \dot{\mathbf{x}}\cdot\nabla\Phi\right).$$

Using Eq. (3.9), we readily find $m\mathbf{v}\cdot\dot{\mathbf{v}} = e\,\mathbf{E}\cdot\mathbf{v}$ and thus

$$-\,e\,\mathbf{E}\cdot\mathbf{v} \;-\; e\left(\frac{\partial\Phi}{\partial t} \;+\; \dot{\mathbf{x}}\cdot\nabla\Phi\right) \;=\; \frac{e}{c}\,\frac{\partial\mathbf{A}}{\partial t}\cdot\dot{\mathbf{x}} \;-\; e\,\frac{\partial\Phi}{\partial t},$$

which is shown to be satisfied exactly by substituting the definition for $\mathbf{E}$.

3.4.3 Gauge Invariance

The electric and magnetic fields defined in (3.12) are invariant under the gauge transformation

$$\Phi \;\rightarrow\; \Phi \;-\; \frac{1}{c}\,\frac{\partial \chi}{\partial t} \;\text{ and }\; \mathbf{A} \;\rightarrow\; \mathbf{A} \;+\; \nabla\chi, \tag{3.15}$$

where $\chi(\mathbf{x}, t)$ is an arbitrary scalar field.

Although the equations of motion (3.9) and (3.10) are *manifestly* gauge invariant, the Lagrangian (3.11) is not manifestly gauge invariant since the electromagnetic potentials Φ and $\mathbf{A}$ appear explicitly. Under a gauge transformation, however, we find

$$L \;\;\rightarrow\;\; L \;+\; \frac{e}{c}\,\nabla\chi\cdot\dot{\mathbf{x}} \;-\; e\left(-\,\frac{1}{c}\,\frac{\partial\chi}{\partial t}\right) \;=\; L \;+\; \frac{e}{c}\,\frac{d\chi}{dt}.$$

As is generally known, since Lagrangians can only be defined up to the exact time derivative of a time-dependent function on configuration space (i.e., *equivalent* Lagrangians yield the same Euler-Lagrange equations), we find that a gauge transformation keeps the Lagrangian within the same equivalence class.

3.4.4 Canonical Hamilton's Equationss

The canonical momentum $\mathbf{p}$ for a particle of mass m and charge e in an electromagnetic field is defined as

$$\mathbf{p}(\mathbf{x}, \mathbf{v}, t) \;=\; \frac{\partial L}{\partial \dot{\mathbf{x}}} \;=\; m\,\mathbf{v} \;+\; \frac{e}{c}\,\mathbf{A}(\mathbf{x}, t). \tag{3.16}$$

The canonical Hamiltonian function $H(\mathbf{x}, \mathbf{p}, t)$ is now constructed through the Legendre transformation

$$\begin{aligned} H(\mathbf{x}, \mathbf{p}, t) \;&=\; \mathbf{p}\cdot\dot{\mathbf{x}}(\mathbf{x}, \mathbf{p}, t) \;-\; L[\mathbf{x}, \dot{\mathbf{x}}(\mathbf{x}, \mathbf{p}, t), t] \\ &=\; e\,\Phi(\mathbf{x}, t) \;+\; \frac{1}{2m}\left|\,\mathbf{p} \;-\; \frac{e}{c}\,\mathbf{A}(\mathbf{x}, t)\,\right|^2, \end{aligned} \tag{3.17}$$

where $\mathbf{v}(\mathbf{x}, \mathbf{p}, t)$ was obtained by inverting $\mathbf{p}(\mathbf{x}, \mathbf{v}, t)$ from Eq. (3.16). Using the canonical Hamiltonian function (3.17), we immediately find

$$\begin{aligned} \dot{\mathbf{x}} \;&=\; \frac{\partial H}{\partial \mathbf{p}} \;=\; \frac{1}{m}\left(\mathbf{p} \;-\; \frac{e}{c}\,\mathbf{A}\right), \\ \dot{\mathbf{p}} \;&=\; -\,\frac{\partial H}{\partial \mathbf{x}} \;=\; -\,e\,\nabla\Phi \;-\; \frac{e}{c}\,\nabla\mathbf{A}\cdot\dot{\mathbf{x}}, \end{aligned}$$

from which we recover the equations of motion (3.9) and (3.10) once we use the definition (3.16) for the canonical momentum.

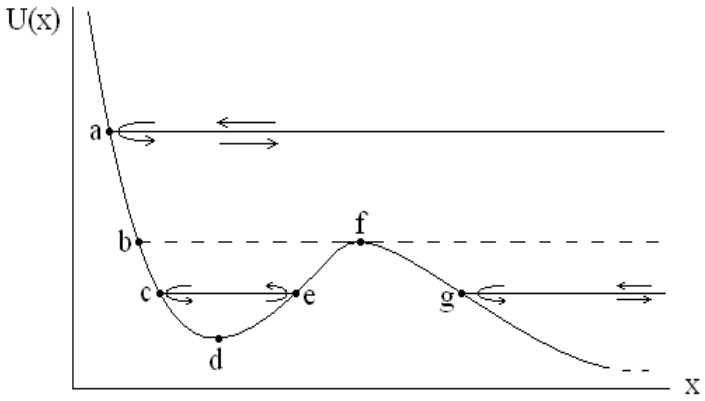


Figure 3.1: Bounded and unbounded orbits

3.5 One-degree-of-freedom Hamiltonian Dynamics

The one degree-of-freedom Hamiltonian dynamics of a particle of mass m is based on the Hamiltonian

$$H(x,p) \;=\; \frac{p^2}{2m} \;+\; U(x), \tag{3.18}$$

where $p = m\dot{x}$ is the particle's momentum and $U(x)$ is the potential energy. The Hamilton's equations (3.2) for this Hamiltonian are

$$\frac{dx}{dt} \;=\; \frac{p}{m} \quad \text{and} \quad \frac{dp}{dt} \;=\; -\,\frac{dU(x)}{dx}. \tag{3.19}$$

Since the Hamiltonian (and Lagrangian) is time independent, the energy conservation law states that $H(x,p) = E$. In turn, this conservation law implies that the particle's velocity $\dot{x}$ can be expressed as

$$\dot{x}(x,E) \;=\; \pm\sqrt{\frac{2}{m}\,[E - U(x)]}, \tag{3.20}$$

where the sign of $\dot{x}$ is determined from the initial conditions.

It is immediately clear that physical motion is possible only if $E \geq U(x)$; points where $E = U(x)$ are known as *turning* points. In Figure 3.1, each horizontal line corresponds to a constant energy value (called an energy *level*). For the top energy level, only one turning point (labeled a in Figure 3.1) exists and a particle coming from the right will be *reflected* at point a and return to large values of x; the motion in this case is said to be *unbounded*. As the energy value is lowered, two turning points (labeled b and f) exist and motion can either be *bounded* (between points b and f) or unbounded (if the initial position is to the right of point f); this energy level is known as the *separatrix* level since bounded and unbounded motions share one turning point. As energy is lowered below the separatrix

level, three turning points (labeled c, e, and g) exist and, once again, motion can either be *bounded* (between points c and e) or unbounded (if the initial position is to the right of point g).[1] Lastly, we note that point d in Figure 3.1 is actually an equilibrium point (as is point f); only unbounded motion is allowed as energy is lowered below point d.

The dynamical solution $x(t;E)$ of the Hamilton's equations (3.19) is first expressed an integration by quadrature as

$$t(x;E) \;=\; \sqrt{\frac{m}{2}} \int_{x_0}^{x} \frac{ds}{\sqrt{E-U(s)}}, \tag{3.21}$$

where x_0 is the particle's initial position is between turning $x_1 < x_2$ (allowing $x_2 \to \infty$) and we assume that $\dot{x}(0) > 0$. Next, inversion of the relation (3.21) yields the solution $x(t;E)$.

For bounded motion in one dimension, the particle bounces back and forth between the two turning points x_1 and $x_2 > x_1$, and the period of oscillation $T(E)$ is a function of energy alone

$$T(E) \;=\; 2 \int_{x_1}^{x_2} \frac{dx}{\dot{x}(x,E)} \;=\; \sqrt{2m} \int_{x_1}^{x_2} \frac{dx}{\sqrt{E-U(x)}}. \tag{3.22}$$

3.5.1 Simple Harmonic Oscillator

As a first example, we consider the case of a particle of mass m attached to a spring of constant k, for which the potential energy is $U(x) = \frac{1}{2}\,kx^2$. The motion of a particle with total energy E is always bounded, with turning points

$$x_{1,2} \;=\; \pm\sqrt{2E/k} \;=\; \pm\, a.$$

We start with the solution $t(x;E)$ for the case of $x(0;E) = +a$, so that $\dot{x}(t;E) < 0$ for $t > 0$, and

$$t(x;E) \;=\; \sqrt{\frac{m}{k}} \int_{x}^{a} \frac{ds}{\sqrt{a^2-s^2}} \;=\; \sqrt{\frac{m}{k}}\, \arccos\left(\frac{x}{a}\right). \tag{3.23}$$

Inversion of this relation yields the well-known solution $x(t;E) = a\,\cos(\omega_0 t)$, where $\omega_0 = \sqrt{k/m}$. Using Eq. (3.22), we find the period of oscillation

$$T(E) \;=\; \frac{4}{\omega_0} \int_{0}^{a} \frac{dx}{\sqrt{a^2-x^2}} \;=\; \frac{2\pi}{\omega_0},$$

which turns out to be independent of energy E.

[1] Note: Quantum tunneling involves a connection between the bounded and unbounded solutions.

3.5.2 Pendulum

Our second example involves the case of the pendulum of length ℓ and mass m in a gravitational field g. The Hamiltonian in this case is

$$H \;=\; \frac{1}{2}\, m\ell^2\, \dot\theta^2 \;+\; mg\ell\,(1-\cos\theta).$$

The total energy of the pendulum is determined from its initial conditions $(\theta_0, \dot\theta_0)$:

$$E \;=\; \frac{1}{2}\, m\ell^2\, \dot\theta_0^2 \;+\; mg\ell\,(1-\cos\theta_0),$$

and thus solutions of the pendulum problem are divided into three classes depending on the value of the total energy of the pendulum: Class I (rotation) $E > 2\,mg\ell$, Class II (separatrix) $E = 2\,mg\ell$, and Class III (libration) $E < 2\,mg\ell$.

In the rotation class ($E > 2\,mg\ell$), the kinetic energy can never vanish and the pendulum keeps rotating either clockwise or counter-clockwise depending on the sign of $\dot\theta_0$. In the libration class ($E < 2\,mg\ell$), on the other hand, the kinetic energy vanishes at turning points easily determined by initial conditions if the pendulum starts from rest – in this case, the turning points are $\pm\,\theta_0$, where

$$\theta_0 \;=\; \arccos\left(1 \;-\; \frac{E}{mg\ell}\right).$$

In the separatrix class ($E = 2\,mg\ell$), the turning points are $\pm\,\pi$. The numerical solution of the normalized pendulum equation $\ddot\theta + \sin\theta = 0$ subject to the initial condition θ_0 and $\dot\theta_0 = \pm\sqrt{2\,(\epsilon - 1 + \cos\theta_0)}$ yields the following curves. Here, the three classes I, II, and III are easily seen (with $\epsilon = 1 - \cos\theta_0$ and $\dot\theta_0 = 0$ for classes I and II and $\epsilon > 1 - \cos\theta_0$ for class III). Note that for rotations (class III), the pendulum slows down as it approaches $\theta = \pm\,\pi$ (the top part of the circle) and speeds up as it approaches $\theta = 0$ (the bottom part of the circle). In fact, since $\theta = \pi$ and $\theta = -\pi$ represent the same point in space, the lines AB and $A'B'$ in Figure 3.2 should be viewed as being identical (i.e., they should be glued together) and the geometry of the *phase space* for the pendulum problem is actually that of a cylinder.

We now look at an explicit solution for pendulum librations (class I), where the angular velocity $\dot\theta$ is

$$\dot\theta(\theta; E) \;=\; \pm\,\omega_0\,\sqrt{2\,(\cos\theta - \cos\theta_0)} \;=\; \pm\,2\omega_0\,\sqrt{\sin^2(\theta_0/2) \;-\; \sin^2(\theta/2)}, \qquad (3.24)$$

where $\omega_0 = \sqrt{g/\ell}$ denotes the characteristic angular frequency and, thus, $\pm\,\theta_0$ are the turning points for this problem. By making the substitution $\sin\theta/2 = k\,\sin\varphi$, where

$$k(E) \;=\; \sin[\,\theta_0(E)/2] \;=\; \sqrt{\frac{E}{2\,mg\ell}} \;<\; 1$$

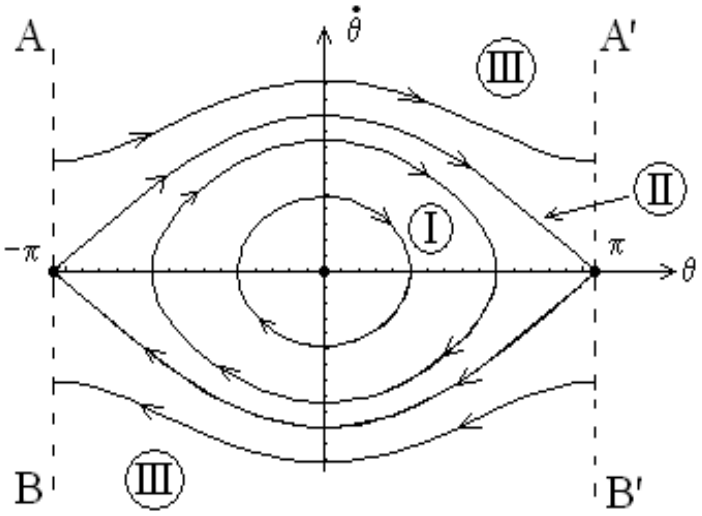


Figure 3.2: Phase space of the pendulum

and $\varphi = \pm\,\pi/2$ when $\theta = \pm\,\theta_0$, the libration solution of the pendulum problem is thus

$$\omega_0\, t(\theta; E) \;=\; \int_{\Theta(\theta;E)}^{\pi/2} \frac{d\varphi}{\sqrt{1 - k^2\,\sin^2\varphi}}, \tag{3.25}$$

where $\Theta(\theta; E) = \arcsin(k^{-1}\sin\theta/2)$. The inversion of this relation yields $\theta(t; E)$ expressed in terms of elliptic functions, while the period of oscillation is defined as

$$\begin{aligned} \omega_0\, T(E) \;&=\; 4\int_0^{\pi/2} \frac{d\varphi}{\sqrt{1 - k^2\,\sin^2\varphi}} \;=\; 4\int_0^{\pi/2} d\varphi \left(1 \,+\, \frac{k^2}{2}\,\sin^2\varphi \,+\, \cdots\right), \\ &=\; 2\pi\left(1 \,+\, \frac{k^2}{4} \,+\, \cdots\right) \;=\; 4\,K(k^2), \end{aligned} \tag{3.26}$$

where $K(k^2)$ denotes the complete elliptic integral of the first kind (see Figure 3.3). We note here that if $k \ll 1$ (or $\theta_0 \ll 1$ rad) the libration period of a pendulum is nearly independent of energy, $T \simeq 2\pi/\omega_0$. However, we also note that as $E \to 2\,mg\ell$ ($k \to 1$ or $\theta_0 \to \pi$ rad), the libration period of the pendulum becomes infinitely large (see Figure above), i.e., $T \to \infty$ as $k \to 1$.

In this separatrix limit ($\theta_0 \to \pi$), the pendulum equation (3.24) yields the *separatrix* equation $\dot\varphi = \omega_0\,\cos\varphi$, where $\varphi = \theta/2$. The separatrix solution is expressed in terms of the transcendental equation

$$\sec\varphi(t) \;=\; \cosh(\omega_0 t + \gamma),$$

where $\cosh\gamma = \sec\varphi_0$ represents the initial condition. We again note that $\varphi \to \pi/2$ (or $\theta \to \pi$) only as $t \to \infty$. Separatrices are quite common in periodic dynamical systems as will be shown in Sec. 3.6 and Sec. 7.2.3.

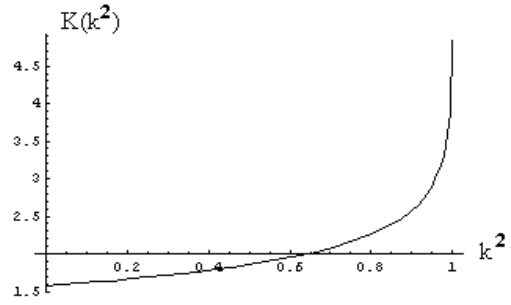


Figure 3.3: Pendulum period as a function of energy

3.5.3 Constrained Motion on the Surface of a Cone

The constrained motion of a particle of mass m on a cone in the presence of gravity was shown in Sec. 2.4.4 to be doubly periodic in the generalized coordinates s and θ. The fact that the Lagrangian (2.20) is independent of time leads to the conservation law of energy

$$E \;=\; \frac{m}{2}\,\dot{s}^2 \;+\; \left(\frac{\ell^2}{2m\sin^2\alpha\, s^2} \;+\; mg\cos\alpha\, s\right) \;=\; \frac{m}{2}\,\dot{s}^2 \;+\; V(s), \tag{3.27}$$

where we have taken into account the conservation law of angular momentum $\ell = ms^2\sin^2\alpha\,\dot{\theta}$. The effective potential $V(s)$ has a single minimum $V_0 = \frac{3}{2}\,mgs_0\,\cos\alpha$ at

$$s_0 \;=\; \left(\frac{\ell^2}{m^2 g\;\sin^2\alpha\;\cos\alpha}\right)^{\frac{1}{3}},$$

and the only type of motion is bounded when $E > V_0$. The turning points for this problem are solutions of the cubic equation

$$\frac{3}{2}\,\epsilon \;=\; \frac{1}{2\sigma^2} \;+\; \sigma,$$

where $\epsilon = E/V_0$ and $\sigma = s/s_0$. Figure 3.4 shows the evolution of the three roots of this equation as the normalized energy parameter ϵ is varied. The three roots $(\sigma_0, \sigma_1, \sigma_2)$ satisfy the relations $\sigma_0\sigma_1\sigma_2 = -\frac{1}{2}$, $\sigma_0+\sigma_1+\sigma_2 = \frac{3}{2}\,\epsilon$, and $\sigma_1^{-1}+\sigma_2^{-1} = -\,\sigma_0^{-1}$. We see that one root (labeled σ_0) remains negative for all normalized energies ϵ; this root is unphysical since s must be positive (by definition). On the other hand, the other two roots (σ_1, σ_2), which are complex for $\epsilon < 1$ (i.e., for energies below the minumum of the effective potential energy V_0), become real at $\epsilon = 1$, where $\sigma_1 = \sigma_2$, and separate $(\sigma_1 < \sigma_2)$ for larger values of ϵ (in the limit $\epsilon \gg 1$, we find $\sigma_2 \simeq \frac{3}{2}\,\epsilon$ and $\sigma_1^{-1} \simeq -\,\sigma_0^{-1} \simeq \sqrt{3\epsilon}$). Lastly, the period of oscillation is determined by the definite integral

$$T(\epsilon) \;=\; 2\,\sqrt{\frac{s_0}{g\;\cos\alpha}}\;\int_{\sigma_1}^{\sigma_2}\;\frac{\sigma\;d\sigma}{\sqrt{3\epsilon\,\sigma^2\;-\;1\;-\;2\,\sigma^3}},$$

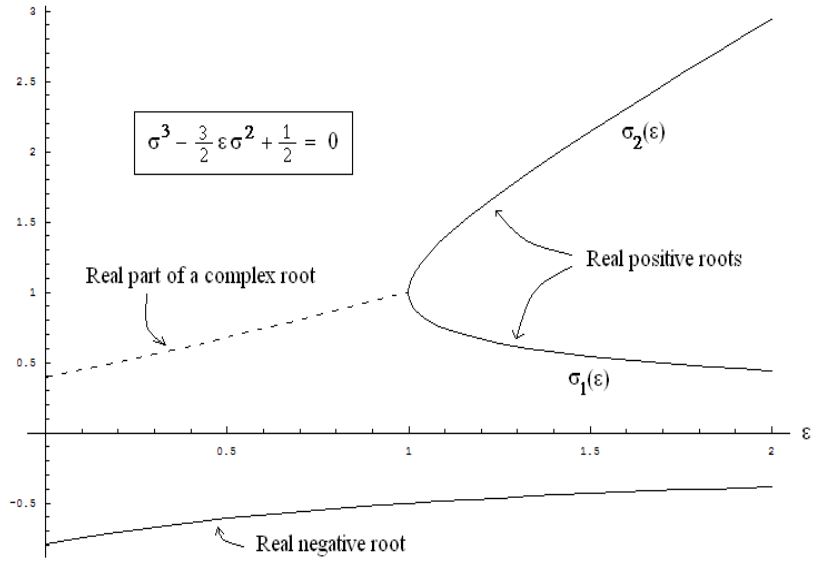


Figure 3.4: Roots of a cubic equation

whose solution is expressed in terms of elliptic integrals.

3.6 Charged Spherical Pendulum in a Magnetic Field*

A spherical pendulum of length ℓ and mass m carries a positive charge e and moves under the action of a constant gravitational field (with acceleration g) and a constant magnetic field B (see Figure 3.5). The position vector of the pendulum is

$$\mathbf{x} \;=\; \ell\,[\,\sin\theta\;(\cos\varphi\,\widehat{x} + \sin\varphi\,\widehat{y}) \;-\; \cos\theta\,\widehat{z}\,]\,,$$

and thus its velocity $\mathbf{v} = \dot{\mathbf{x}}$ is

$$\mathbf{v} \;=\; \ell\,\dot{\theta}\;[\,\cos\theta\;(\cos\varphi\,\widehat{x} + \sin\varphi\,\widehat{y}) \;+\; \sin\theta\,\widehat{z}\,] \;+\; \ell\,\sin\theta\;\dot{\varphi}\;(-\,\sin\varphi\,\widehat{x} + \cos\varphi\,\widehat{y})\,,$$

and the kinetic energy of the pendulum is

$$K \;=\; \frac{m\,\ell^2}{2}\left(\dot{\theta}^2 \;+\; \sin^2\theta\;\dot{\varphi}^2\right).$$

3.6.1 Lagrangian

Because the charged pendulum moves in a magnetic field $\mathbf{B} = -\,B\,\widehat{z}$, we must include the magnetic term $\mathbf{v}\cdot e\mathbf{A}/c$ in the Lagrangian [see Eq. (3.11)]. Here, the vector potential $\mathbf{A}$

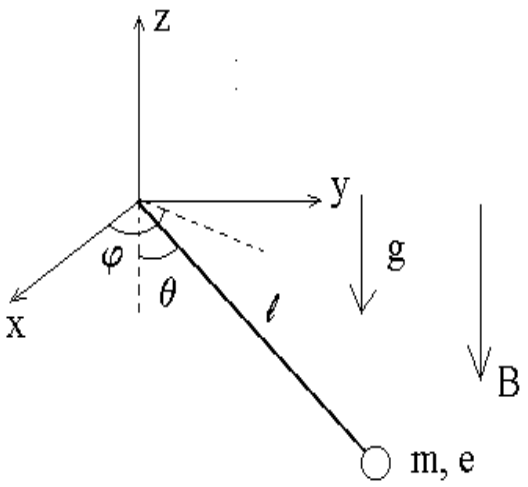


Figure 3.5: Charged pendulum in a magnetic field

must be evaluated at the position of the pendulum and is thus expressed as

$$\mathbf{A} \;=\; -\,\frac{B\ell}{2}\,\sin\theta\,(-\,\sin\varphi\,\widehat{x} + \cos\varphi\,\widehat{y})\,,$$

and, hence, we find

$$\frac{e}{c}\,\mathbf{A}\cdot\mathbf{v} \;=\; -\,\frac{B\ell^2}{2}\,\sin^2\theta\,\dot{\varphi}.$$

Lastly, the charged pendulum is under the influence of two potential energy terms: gravitational potential energy $mg\ell\,(1-\cos\theta)$ and magnetic potential energy $(-\boldsymbol{\mu}\cdot\mathbf{B} = \mu_z\,B)$, where

$$\boldsymbol{\mu} \;=\; \frac{e}{2mc}\,(\mathbf{x}\times\mathbf{v})$$

denotes the magnetic moment of a charge e moving about a magnetic field line. Here, it is easy to find

$$\mu_z\;B \;=\; -\,\frac{eB}{2c}\,\ell^2\sin^2\theta\,\dot{\varphi}$$

and by combining the various terms, the Lagrangian for the system is

$$L(\theta,\dot{\theta},\;\dot{\varphi}) \;=\; m\ell^2\left[\;\frac{\dot{\theta}^2}{2}\;+\;\sin^2\theta\left(\frac{\dot{\varphi}^2}{2}\;-\;\omega_c\,\dot{\varphi}\right)\right]\;-\;mg\ell\,(1-\cos\theta), \tag{3.28}$$

where the *cyclotron* frequency ω_c is defined as $\omega_c = eB/mc$.

3.6.2 Euler-Lagrange equations

The Euler-Lagrange equation for θ is

$$\frac{\partial L}{\partial \dot{\theta}} \;=\; m\ell^2\,\dot{\theta} \;\;\rightarrow\;\; \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) \;=\; m\ell^2\,\ddot{\theta}$$

$$\frac{\partial L}{\partial \theta} \;=\; -\,mg\ell\,\sin\theta \;+\; m\ell^2\,\left(\dot{\varphi}^2 - 2\,\omega_c\,\dot{\varphi}\right)\,\sin\theta\,\cos\theta$$

or

$$\ddot{\theta} \;+\; \frac{g}{\ell}\,\sin\theta \;=\; \left(\dot{\varphi}^2 - 2\,\omega_c\,\dot{\varphi}\right)\,\sin\theta\,\cos\theta$$

The Euler-Lagrange equation for φ immediately leads to a constant of the motion for the system since the Lagrangian (3.28) is independent of the azimuthal angle φ and hence

$$p_\varphi \;=\; \frac{\partial L}{\partial \dot{\varphi}} \;=\; m\ell^2\,\sin^2\theta\;\left(\dot{\varphi} - \omega_c\right)$$

is a constant of the motion, i.e., the Euler-Lagrange equation for φ states that $\dot{p}_\varphi = 0$.

Since p_φ is a constant of the motion, we can use it the rewrite $\dot{\varphi}$ in the Euler-Lagrange equation for θ as

$$\dot{\varphi} \;=\; \omega_c \;+\; \frac{p_\varphi}{m\ell^2\,\sin^2\theta},$$

and thus

$$\dot{\varphi}^2 \;-\; 2\,\omega_c\,\dot{\varphi} \;=\; \left(\frac{p_\varphi}{m\ell^2\,\sin^2\theta}\right)^2 \;-\; \omega_c^2$$

so that

$$\ddot{\theta} \;+\; \frac{g}{\ell}\,\sin\theta \;=\; \sin\theta\,\cos\theta\left[\;\left(\frac{p_\varphi}{m\ell^2\,\sin^2\theta}\right)^2 \;-\; \omega_c^2\;\right]. \tag{3.29}$$

Not surprisingly the integration of the second-order differential equation for θ is complex (see below). It turns out, however, that the Hamiltonian formalism allows us glimpses into the gobal structure of general solutions of this equation. Lastly, we note that the Routh-Lagrange function $R(\theta, \dot{\theta}; p_\varphi)$ for this problem is

$$R(\theta, \dot{\theta}; p_\varphi) \;=\; L \;-\; \dot{\varphi}\,\frac{\partial L}{\partial \dot{\varphi}} \;=\; \frac{m\ell^2}{2}\,\dot{\theta}^2 \;-\; V(\theta; p_\varphi),$$

where

$$V(\theta; p_\varphi) \;=\; mg\ell\,(1 - \cos\theta) \;+\; \frac{1}{2m\ell^2\,\sin^2\theta}\left(p_\varphi \;+\; m\ell^2\omega_c\,\sin^2\theta\right)^2 \tag{3.30}$$

represents an *effective* potential under which the charged spherical pendulum moves, so that the Euler-Lagrange equation (3.29) may be written as

$$\frac{d}{dt}\left(\frac{\partial R}{\partial \dot{\theta}}\right) \;=\; \frac{\partial R}{\partial \theta} \quad\rightarrow\quad m\ell^2\,\ddot{\theta} \;=\; -\,\frac{\partial V}{\partial \theta}.$$

3.6.3 Hamiltonian

The Hamiltonian for the system is obtained through the Legendre transformation

$$\begin{aligned} H &= \dot\theta \frac{\partial L}{\partial \dot\theta} + \dot\varphi \frac{\partial L}{\partial \dot\varphi} - L \\ &= \frac{p_\theta^2}{2m\ell^2} + \frac{1}{2m\ell^2 \sin^2\theta}\left(p_\varphi + m\ell^2\omega_c \sin^2\theta\right)^2 + mg\ell\,(1-\cos\theta). \end{aligned} \tag{3.31}$$

The Hamilton's equations for (θ, p_θ) are

$$\begin{aligned} \dot\theta &= \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{m\ell^2} \\ \dot p_\theta &= -\frac{\partial H}{\partial \theta} = -mg\ell \sin\theta + m\ell^2 \sin\theta\cos\theta\left[\left(\frac{p_\varphi}{m\ell^2\sin^2\theta}\right)^2 - \omega_c^2\right], \end{aligned}$$

while the Hamilton's equations for (φ, p_φ) are

$$\begin{aligned} \dot\varphi &= \{\varphi,\, H\} = \frac{\partial H}{\partial p_\varphi} = \frac{p_\varphi}{m\ell^2\sin^2\theta} + \omega_c \\ \dot p_\varphi &= \{p_\varphi,\, H\} = -\frac{\partial H}{\partial \varphi} = 0. \end{aligned}$$

It is readily checked that these Hamilton equations lead to the same equations as the Euler-Lagrange equations for θ and φ.

So what have we gained? It turns out that a most useful application of the Hamiltonian formalism resides in the use of the constants of the motion to plot Hamiltonian *orbits* in phase space. Indeed, for the problem considered here, a Hamiltonian orbit is expressed in the form $p_\theta(\theta;\ E, p_\varphi)$, i.e., each orbit is labeled by values of the two constants of motion E (the total energy) and p_ϕ the azimuthal canonical momentum (actually an angular momentum):

$$p_\theta = \pm\sqrt{2m\ell^2\,(E - mg\ell + mg\ell\cos\theta) - \frac{1}{\sin^2\theta}\left(p_\varphi + m\ell^2\sin^2\theta\,\omega_c\right)^2}.$$

Hence, for charged pendulum of given mass m and charge e with a given cyclotron frequency ω_c (and g), we can completely determine the motion of the system once initial conditions are known (from which E and p_φ can be calculated).

By using the following dimensionless parameters $\Omega = \omega_c/\omega_g$ and $\alpha = p_\varphi/(m\ell^2\omega_g)$, we may write the effective potential (3.30) in dimensionless form $\overline{V}(\theta) = V(\theta)/(mg\ell)$ as

$$\overline{V}(\theta) = 1 - \cos\theta + \frac{1}{2}\left(\frac{\alpha}{\sin\theta} + \Omega\sin\theta\right)^2.$$

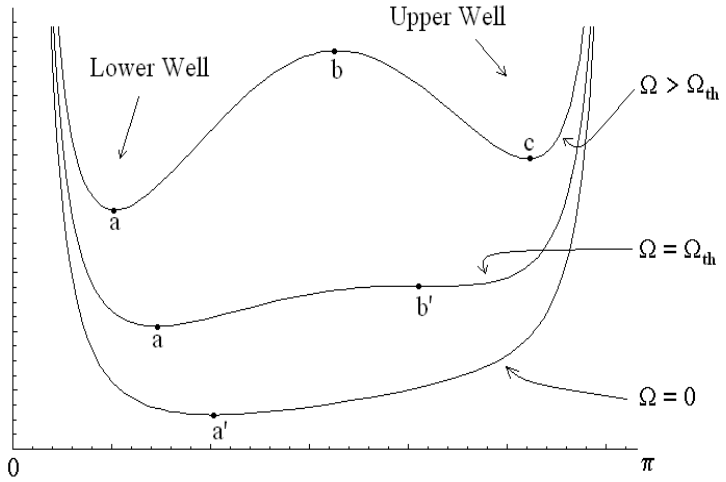


Figure 3.6: Effective potential of the charged pendulum in a magnetic field

Figure 3.6 shows the dimensionless effective potential $\overline{V}(\theta)$ for $\alpha = 1$ and several values of the dimensionless parameter Ω. When Ω is below the *threshold* value $\Omega_{th} = 1.94204...$ (for $\alpha = 1$), the effective potential has a single local minimum (point a' in Figure 3.6). At threshold ($\Omega = \Omega_{th}$), an inflection point develops at point b'. Above this threshold ($\Omega > \Omega_{th}$), a local maximum (at point b) and two local minima (at points a and c) appear. Note that the local maximum at point b implies the existence of a separatrix solution, which separates the bounded motion in the lower well and the upper well.

The normalized Euler-Lagrange equations are

$$\varphi' \;=\; \Omega \;+\; \frac{\alpha}{\sin^2\theta} \quad \text{and} \quad \theta'' \;+\; \sin\theta \;=\; \sin\theta\,\cos\theta\left(\frac{\alpha^2}{\sin^4\theta} \;-\; \Omega^2\right)$$

where $\tau = \omega_g t$ denotes the dimensionless time parameter and the dimensionless parameters are defined in terms of physical constants.

Figures 3.7 below show three-dimensional spherical projections (first row) and (x, y)-plane projections (second row) for three cases above threshold ($\Omega > \Omega_{th}$): motion in the lower well (left column), separatrix motion (center column), and motion in the upper well (right column). Figure 3.8 shows the (x, z)-plane projections for these three cases combined on the same graph. These Figures clearly show that a separatrix solution exists which separates motion in either the upper well or the lower well.

The equation for φ' does not change sign if $\alpha > -\Omega$, while its sign can change if $\alpha < -\Omega$ (or $p_\varphi < -m\ell^2\,\omega_c$). Figures 3.9 show the effect of changing $\alpha \to -\alpha$ by showing the graphs θ versus φ (first row), the (x, y)-plane projections (second row), and the (x, z)-plane projections (third row). One can clearly observe the wonderfully complex dynamics

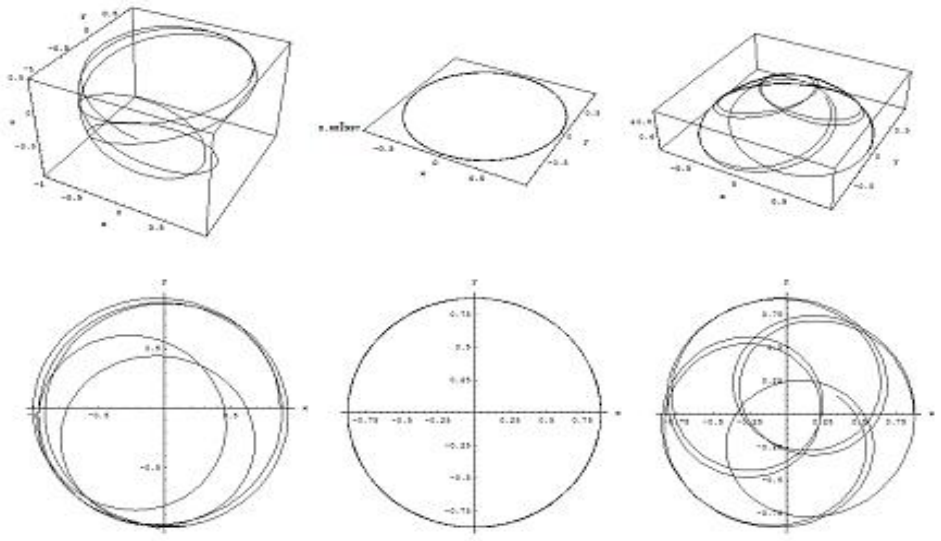


Figure 3.7: Orbits of the charged pendulum

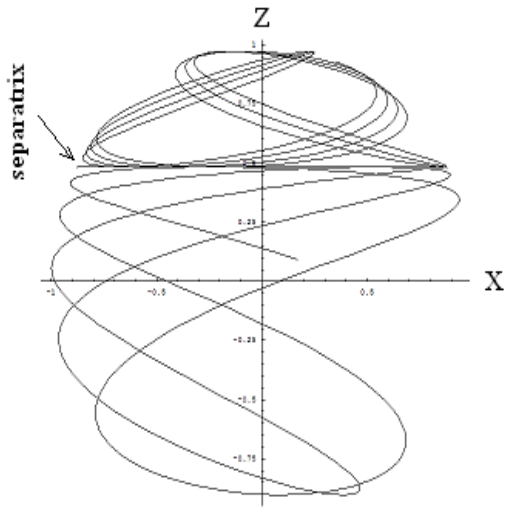


Figure 3.8: Orbit projections

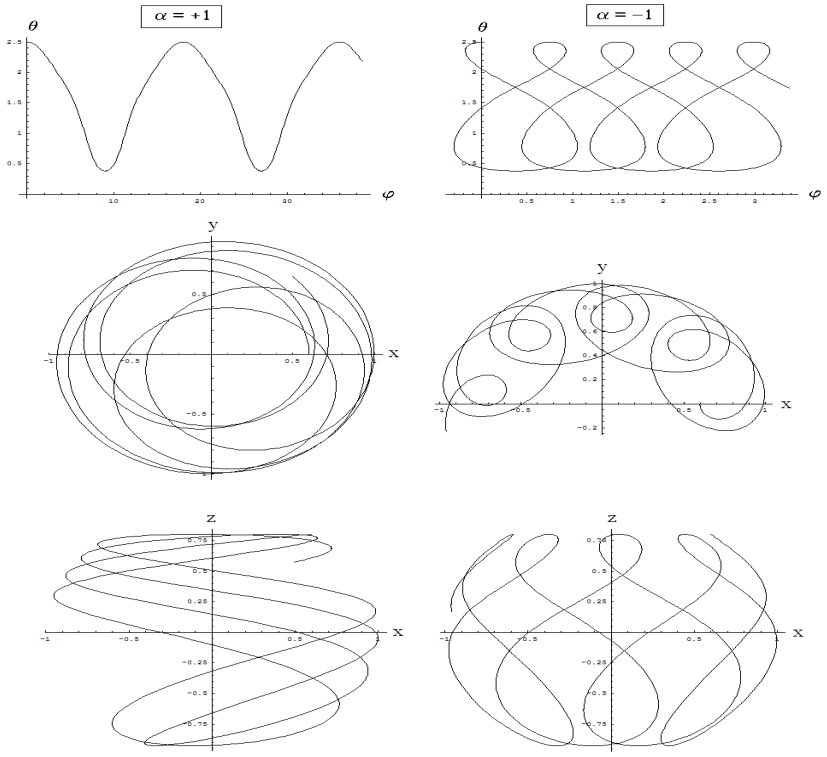


Figure 3.9: Retrograde motion

of the charged pendulum in a uniform magnetic field, which is explicitly characterized by the effective potential $V(\theta)$ given by Eq. (3.30).

3.7 Problems

Problem 1

A particle of mass m and total energy E moves periodically in a one-dimensional potential given as

$$U(x) \;=\; F\,|x| \;=\; \begin{cases} F\,x & (x > 0) \\ \\ -\,F\,x & (x < 0) \end{cases}$$

where F is a positive constant.

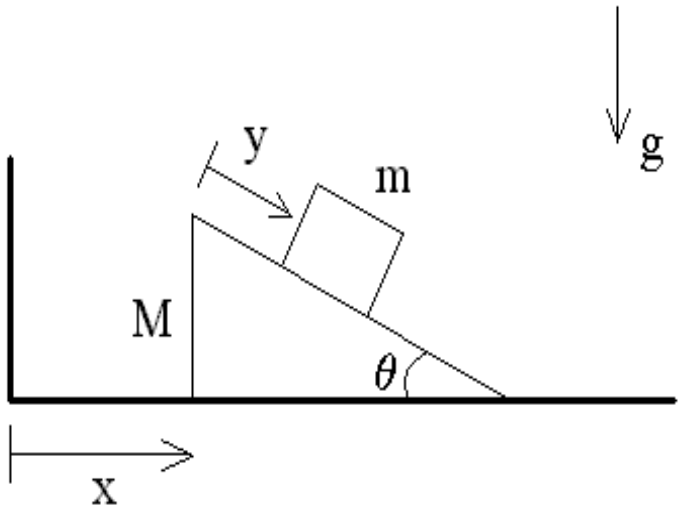
(a) Find the turning points for this potential.

(b) Find the dynamical solution $x(t; E)$ for this potential by choosing a suitable initial condition.

(c) Find the period $T(E)$ for the motion.

Problem 2

A block of mass m rests on the inclined plane (with angle θ) of a triangular block of mass M as shown in the Figure below. Here, we consider the case where both blocks slide without friction (i.e., m slides on the inclined plane without friction and M slides without friction on the horizontal plane).



(a) Using the generalized coordinates (x, y) shown in the Figure above, construct the La-

grangian $L(x, \dot{x},\ y, \dot{y})$.

(b) Derive the Euler-Lagrange equations for x and y.

(c) Calculate the canonical momenta

$$p_x(x, \dot{x},\ y, \dot{y}) \ = \ \frac{\partial L}{\partial \dot{x}} \quad \text{and} \quad p_y(x, \dot{x},\ y, \dot{y}) \ = \ \frac{\partial L}{\partial \dot{y}},$$

and invert these expressions to find the functions $\dot{x}(x, p_x,\ y, p_y)$ and $\dot{y}(x, p_x,\ y, p_y)$.

(d) Calculate the Hamiltonian $H(x, p_x,\ y, p_y)$ for this system by using the Legendre transformation

$$H(x, p_x,\ y, p_y) \ = \ p_x\, \dot{x} \ + \ p_y\, \dot{y} \ - \ L(x, \dot{x},\ y, \dot{y}),$$

where the functions $\dot{x}(x, p_x,\ y, p_y)$ and $\dot{y}(x, p_x,\ y, p_y)$ are used.

(e) Find which of the two momenta found in Part (c) is a constant of the motion and discuss why it is so. If the two blocks start from rest, what is the value of this constant of motion?

Chapter 4

Motion in a Central-Force Field

4.1 Motion in a Central-Force Field

A particle moves under the influence of a central-force field $\mathbf{F}(\mathbf{r}) = F(r)\,\hat{\mathbf{r}}$ if the force on the particle is independent of the angular position of the particle about the center of force and depends only on its distance r from the center of force. Here, the magnitude $F(r)$ (which is positive for a repulsive force and negative for an attractive force) is defined in terms of the central potential $U(r)$ as $F(r) = -\,U'(r)$. Note that for a central-force potential $U(r)$, the angular momentum $\mathbf{L}$ in the CM frame is a constant of the motion since $\mathbf{r} \times \nabla U(r) = 0$.

4.1.1 Lagrangian Formalism

The motion of two particles in an isolated system takes place on a two-dimensional plane; we, henceforth, this plane to be the (x, y)-plane for which the angular momentum is $\mathbf{L} = \ell\,\hat{\mathbf{z}}$. When these particles move in a central-force field, the Lagrangian is simply

$$L \;=\; \frac{\mu}{2}\left(\dot{r}^2 \;+\; r^2\,\dot{\theta}^2\right) \;-\; U(r), \tag{4.1}$$

where polar coordinates (r, θ) are most conveniently used, with $x = r\,\cos\theta$ and $y = r\,\sin\theta$. Since the potential U is independent of θ, the canonical angular momentum

$$p_\theta \;=\; \frac{\partial L}{\partial \dot{\theta}} \;=\; \mu\,r^2\,\dot{\theta} \;\equiv\; \ell \tag{4.2}$$

is a constant of motion (here, labeled ℓ). The Euler-Lagrange equation for r, therefore, becomes the radial force equation

$$\mu\left(\ddot{r} \;-\; r\,\dot{\theta}^2\right) \;=\; \mu\,\ddot{r} \;-\; \frac{\ell^2}{\mu\,r^3} \;=\; F(r). \tag{4.3}$$

In this description, the planar orbit is parametrized by time, i.e., once $r(t)$ and $\theta(t)$ are obtained, a path $r(\theta)$ onto the plane is defined.

Since $\dot{\theta}$ does not change sign on its path along the orbit, we may replace $\dot{r}$ and $\ddot{r}$ with $r'(\theta)$ and $r''(\theta)$ as follows. First, we begin with

$$\dot{r} \;=\; \dot{\theta}\, r' \;=\; \frac{\ell\, r'}{\mu\, r^2} \;=\; -\,\frac{\ell}{\mu}\left(\frac{1}{r}\right)' \;=\; -\,(\ell/\mu)\; s',$$

where we use the conservation of angular momentum and define the new dependent variable $s(\theta) = 1/r(\theta)$. Next, we write $\ddot{r} = -\,(\ell/\mu)\,\dot{\theta}\, s'' = -\,(\ell/\mu)^2\, s^2\, s''$, so that the radial force equation (4.3) becomes

$$s'' \;+\; s \;=\; -\,\frac{\mu}{\ell^2\, s^2}\; F(1/s) \;=\; -\,\frac{d\overline{U}(s)}{ds}, \tag{4.4}$$

where $\overline{U}(s) = (\mu/\ell^2)\, U(1/s)$ denotes the normalized central potential expressed as a function of s.

Note that the form of the potential $U(r)$ can be calculated from the solution $s(\theta) = 1/r(\theta)$. For example, consider the particle trajectory described in terms of the solution $r(\theta) = r_0\, \sec(\alpha\,\theta)$, where r_0 and $\alpha > 1$ are constants. The radial equation (4.4) then becomes

$$s'' \;+\; s \;=\; -\,\left(\alpha^2 \;-\; 1\right)\; s \;=\; -\,\frac{d\overline{U}(s)}{ds},$$

and thus

$$\overline{U}(s) \;=\; \frac{1}{2}\;\left(\alpha^2 \;-\; 1\right)\; s^2 \quad\rightarrow\quad U(r) \;=\; \frac{\ell^2}{2\mu\, r^2}\left(\alpha^2 \;-\; 1\right).$$

We note here that, as expected, the central potential is either repulsive for $\alpha > 1$ or attractive for $\alpha < 1$ (see Figure 4.1). Note also that the function $\theta(t)$ is determined from the relation

$$\dot{\theta} \;=\; \frac{\ell}{\mu r^2(\theta)} \quad\rightarrow\quad t(\theta) \;=\; \frac{\mu}{\ell}\;\int_0^{\theta}\; r^2(\phi)\; d\phi.$$

Returning to our example, we find

$$t(\theta) \;=\; \frac{\mu r_0^2}{\alpha\ell}\;\int_0^{\alpha\theta}\; \sec^2\phi\; d\phi \;=\; \frac{\mu r_0^2}{\alpha\ell}\;\tan(\alpha\theta) \quad\rightarrow\quad r(t) \;=\; r_0\,\sqrt{1 \;+\; \left(\frac{\alpha\ell\; t}{\mu r_0^2}\right)^2}$$

and the total energy

$$E \;=\; \frac{\alpha^2\,\ell^2}{2\mu r_0^2},$$

is determined from the initial conditions $r(0) = r_0$ and $\dot{r}(0) = 0$.

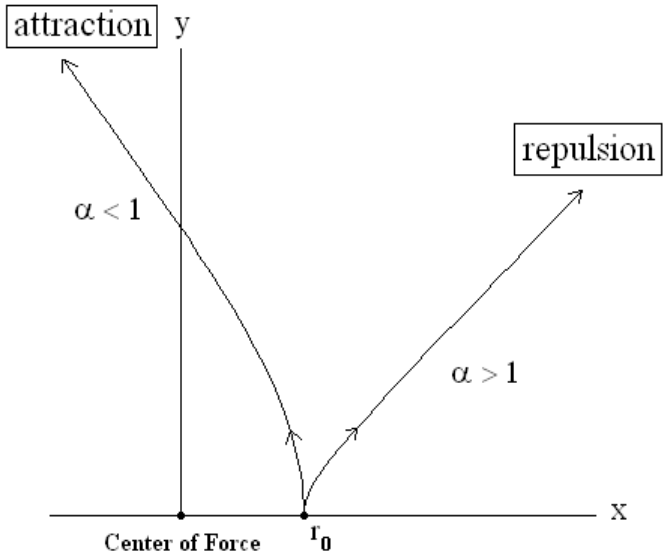


Figure 4.1: Repulsive and attractive orbits

4.1.2 Hamiltonian Formalism

The Hamiltonian for the central-force problem is

$$H \;=\; \frac{p_r^2}{2\mu} \;+\; \frac{\ell^2}{2\,\mu\,r^2} \;+\; U(r),$$

where $p_r = \mu\,\dot{r}$ is the radial canonical momentum and ℓ is the conserved angular momentum. Since energy is also conserved, we solve

$$E \;=\; \frac{\mu\,\dot{r}^2}{2} \;+\; \frac{\ell^2}{2\,\mu\,r^2} \;+\; U(r),$$

as

$$\dot{r} \;=\; \pm\sqrt{\frac{2}{\mu}\,[\,E \;-\; V(r)\,]}, \tag{4.5}$$

where

$$V(r) \;=\; \frac{\ell^2}{2\,\mu\,r^2} \;+\; U(r) \tag{4.6}$$

is known as the *effective* potential and the sign $\pm$ depends on initial conditions. First, we note that this equation yields the integral solution

$$t(r) \;=\; \pm\int \frac{dr}{\sqrt{(2/\mu)\;[E - V(r)]}}. \tag{4.7}$$

This equation can also be used with Eq. (4.2) to yield

$$d\theta \;=\; \frac{\ell}{\mu\,r^2}\,dt \;=\; \frac{\ell}{\mu\,r^2}\,\frac{dr}{\dot{r}} \;=\; \frac{-\,ds}{\sqrt{\epsilon - 2\,\overline{U}(s) \;-\; s^2}}, \tag{4.8}$$

where $\epsilon = 2\mu\, E/\ell^2$, or

$$s'(\theta) \;=\; \pm\sqrt{\epsilon - 2\,\overline{U}(s) \;-\; s^2}. \tag{4.9}$$

We readily check that this equation is a proper solution of the radial force equation (4.4) since

$$s'' \;=\; \frac{s'\,[d\overline{U}/ds + s]}{\sqrt{\epsilon - 2\,\overline{U}(s) \;-\; s^2}} \;=\; -\,\frac{d\overline{U}}{ds} \;-\; s$$

is indeed identical to Eq. (4.4). Hence, for a given central-force potential $U(r)$, we can solve for $r(\theta) = 1/s(\theta)$ by integrating

$$\theta(s) \;=\; -\int_{s_0}^{s} \frac{d\sigma}{\sqrt{\epsilon - 2\,\overline{U}(\sigma) \;-\; \sigma^2}}, \tag{4.10}$$

where s_0 defines $\theta(s_0) = 0$, and performing the inversion $\theta(s) \to s(\theta)$.

4.1.3 Turning Points

Eq. (4.9) yields the following energy equation

$$E \;=\; \frac{\mu}{2}\,\dot{r}^2 \;+\; \frac{\ell^2}{2\mu\, r^2} \;+\; U(r) \;=\; \frac{\ell^2}{2\mu}\left[\;(s')^2 \;+\; s^2 \;+\; 2\,\overline{U}(s)\;\right],$$

where $s' = -\,\mu\dot{r}/\ell = -\,p_r/p_\theta$. *Turning* points are those special values of r_n (or s_n) ($n = 1, 2, ...$) for which

$$E \;=\; U(r_n) \;+\; \frac{\ell^2}{2\mu\, r_n^2} \;=\; \frac{\ell^2}{\mu}\left[\;\overline{U}(s_n) \;+\; \frac{s_n^2}{2}\;\right],$$

i.e., $\dot{r}$ (or s') vanishes at these points. If two non-vanishing turning points $r_2 < r_1 < \infty$ (or $0 < s_1 < s_2$) exist, the motion is said to be *bounded* in the interval $r_2 < r < r_1$ (or $s_1 < s < s_2$), otherwise the motion is *unbounded*. If the motion is bounded, the angular period $\Delta\theta$ is defined as

$$\Delta\theta(s) \;=\; 2\int_{s_1}^{s_2} \frac{ds}{\sqrt{\epsilon - 2\,\overline{U}(s) \;-\; s^2}}. \tag{4.11}$$

Here, the bounded orbit is *closed* only if $\Delta\theta$ is a rational multiple of 2π.

4.2 Homogeneous Central Potentials*

An important class of central potentials is provided by homogeneous potentials that satisfy the condition $U(\lambda\,\mathbf{r}) = \lambda^n\, U(\mathbf{r})$, where λ denotes a rescaling parameter and n denotes the *order* of the homogeneous potential.

4.2.1 The Virial Theorem

The Virial Theorem is an important theorem in Celestial Mechanics and Astrophysics. We begin with the time derivative of the quantity $Q = \sum_i \mathbf{p}_i \cdot \mathbf{r}_i$:

$$\frac{dQ}{dt} \;=\; \sum_i \left(\frac{d\mathbf{p}_i}{dt} \cdot \mathbf{r}_i \;+\; \mathbf{p}_i \cdot \frac{d\mathbf{r}_i}{dt} \right), \tag{4.12}$$

where the summation is over all particles in a mechanical system under the influence of a self-interaction potential

$$U \;=\; \frac{1}{2} \sum_{i,\, j \neq i} U(\mathbf{r}_i - \mathbf{r}_j).$$

We note, however, that since Q itself can be written as a time derivative

$$Q \;=\; \sum_i \, m_i \, \frac{d\mathbf{r}_i}{dt} \cdot \mathbf{r}_i \;=\; \frac{d}{dt} \left(\frac{1}{2} \sum_i \, m_i \, |\mathbf{r}_i|^2 \right) \;=\; \frac{1}{2} \, \frac{d\mathcal{I}}{dt},$$

where $\mathcal{I}$ denotes the *moment of inertia* of the system and that, using Hamilton's equations

$$\frac{d\mathbf{r}_i}{dt} \;=\; \frac{\mathbf{p}_i}{m_i} \quad \text{and} \quad \frac{d\mathbf{p}_i}{dt} \;=\; -\sum_{j \neq i} \nabla_i \, U(\mathbf{r}_i - \mathbf{r}_j),$$

Eq. (4.12) can also be written as

$$\frac{1}{2} \, \frac{d^2\mathcal{I}}{dt^2} \;=\; \sum_i \left(\frac{|\mathbf{p}_i|^2}{m_i} \;-\; \mathbf{r}_i \cdot \sum_{j \neq i} \nabla_i U_{ij} \right) \;=\; 2 \, K \;-\; \sum_{i,\, j \neq i} \mathbf{r}_i \cdot \nabla_i U_{ij}, \tag{4.13}$$

where K denotes the kinetic energy of the mechanical system. Next, using Newton's Third Law, we write

$$\sum_{i,\, j \neq i} \mathbf{r}_i \cdot \nabla_i U_{ij} \;=\; \frac{1}{2} \sum_{i,\, j \neq i} (\mathbf{r}_i - \mathbf{r}_j) \cdot \nabla U(\mathbf{r}_i - \mathbf{r}_j),$$

and, for a homogeneous central potential of order n, we find $\mathbf{r} \nabla U(\mathbf{r}) = n \, U(\mathbf{r})$, so that

$$\frac{1}{2} \sum_{i,\, j \neq i} (\mathbf{r}_i - \mathbf{r}_j) \cdot \nabla U(\mathbf{r}_i - \mathbf{r}_j) \;=\; n \, U.$$

Hence, Eq. (4.13) becomes the *virial of Clausius* (Rudolph Clausius, 1822-1888)

$$\frac{1}{2} \, \frac{d^2\mathcal{I}}{dt^2} \;=\; 2 \, K \;-\; n \, U. \tag{4.14}$$

If we now assume that the mechanical system under consideration is periodic in time, then the time average (denoted $\langle \cdots \rangle$) of Eq. (4.14) yields the Virial Theorem

$$\langle K \rangle \;=\; \frac{n}{2} \, \langle U \rangle, \tag{4.15}$$

so that the time-average of the total energy of the mechanical system, $E = K + U$, is expressed as

$$E \;=\; \begin{cases} (1+n/2)\,\langle U\rangle \\ \\ (1+2/n)\,\langle K\rangle \end{cases}$$

since $\langle E\rangle = E$.

4.2.2 General Properties of Homogeneous Potentials

We now investigate the dynamical properties of orbits in homogeneous central potentials of the form $U(r) = (k/n)\,r^n$ $(n \neq -2)$, where k denotes a positive constant. First, the effective potential (4.6) has an extremum at a distance $r_0 = 1/s_0$ defined as

$$r_0^{n+2} \;=\; \frac{\ell^2}{k\mu} \;=\; \frac{1}{s_0^{n+2}}.$$

It is simple to show that this extremum is a maximum if $n < -2$ or a minimum if $n > -2$; we shall, henceforth, focus our attention on the latter case, where the minimum in the effective potential is

$$V_0 \;=\; V(r_0) \;=\; \left(1 \;+\; \frac{n}{2}\right)\frac{k}{n}\,r_0^n \;=\; \left(1 \;+\; \frac{n}{2}\right)U_0.$$

In the vicinity of this minimum, we can certainly find periodic orbits with turning points $(r_2 = 1/s_2 < r_1 = 1/s_1)$ that satisfy the condition $E = V(r)$.

Next, the radial equation (4.4) is written in terms of the potential $\overline{U}(s) = (\mu/\ell^2)\,U(1/s)$ as

$$s'' \;+\; s \;=\; -\,\frac{d\overline{U}}{ds} \;=\; \frac{s_0^{n+2}}{s^{n+1}},$$

where $s_0 = 1/r_0$ and its solution is given as

$$\theta(s) \;=\; \int_s^{s_2} \frac{d\sigma}{\sqrt{\epsilon - (2/n)\,s_0^{n+2}/\sigma^n \;-\; \sigma^2}}, \tag{4.16}$$

where s_2 denotes the upper turning point in the s-coordinate. The solution (4.16) can be expressed in terms of closed analytic expressions obtained by trigonometric substitution only for $n = -1$ or $n = 2$ (when $\epsilon \neq 0$), which we now study in detail.

4.3 Kepler Problem

In this Section, we solve the Kepler problem where the central potential $U(r) = -\,k/r$ is homogeneous with order $n = -1$ and k is a positive constant. The Virial Theorem (4.15),

therefore, implies that periodic solutions of the Kepler problem have negative total energies $E = -\langle K \rangle = (1/2)\langle U \rangle$.

We now turn to the general solution of the Kepler problem

$$\mu \ddot{r} \;=\; \frac{\ell^2}{\mu r^3} \;-\; \frac{k}{r^2} \quad \text{and} \quad \dot{\theta} \;=\; \frac{\ell}{\mu r^2},$$

whose orbits are either periodic or aperiodic (see Figure 4.2). To obtain an analytic solution $r(\theta)$ for the Kepler problem, as expressed by the radial force equation (4.4), we use the normalized central potential $\overline{U}(s) = -s_0 s$, where $s_0 = \mu k/\ell^2$, and Eq. (4.4) becomes $s'' + s = s_0$. Next, the turning points for the Kepler problem are solutions of the quadratic equation

$$s^2 \;-\; 2\, s_0 \; s \;-\; \epsilon \;=\; 0,$$

which can be written as $s = s_0 \pm \sqrt{s_0^2 + \epsilon}$

$$s_1 \;=\; s_0\,(1 - \mathsf{e}) \quad \text{and} \quad s_2 \;=\; s_0\,(1 + \mathsf{e}),$$

where

$$\mathsf{e} \;=\; \sqrt{1 + \epsilon/s_0^2} \;=\; \sqrt{1 \;+\; 2\, E\ell^2/\mu k^2}.$$

We clearly see from the Figure 4.2 that the effective potential

$$V(r) \;=\; \frac{\ell^2}{2\mu\, r^2} \;-\; \frac{k}{r}$$

for the Kepler problem has a single minimum at $r_0 = \ell^2/(k\mu)$ and that $V_0 = -k/(2r_0)$. We note that motion is bounded (i.e., orbits are periodic) in regions I and II (see Figure 4.2), where $V_0 \le E < 0$ $(0 \le \mathsf{e} < 1)$, and the motion is unbounded (i.e., orbits are aperiodic) in regions III and IV (see Figure 4.2), where $E \ge 0$ $(\mathsf{e} > 1)$.

4.3.1 Bounded Keplerian Orbits

We will now look at the bounded case ($\epsilon < 0$ or $\mathsf{e} < 1$). We define $\theta(s_2) = 0$, so that for the Kepler problem, Eq. (4.10) becomes

$$\theta(s) \;=\; -\int_{s_0\,(1+\mathsf{e})}^{s} \frac{d\sigma}{\sqrt{s_0^2\mathsf{e}^2 - (\sigma - s_0)^2}}, \tag{4.17}$$

which can easily be integrated by using the identity

$$\frac{-\,dx}{\sqrt{a^2 - x^2}} \;=\; d\left[\arccos\left(\frac{x}{a}\right)\right],$$

so that Eq. (4.17) yields

$$\theta(s) \;=\; \arccos\left(\frac{s - s_0}{s_0\,\mathsf{e}}\right),$$

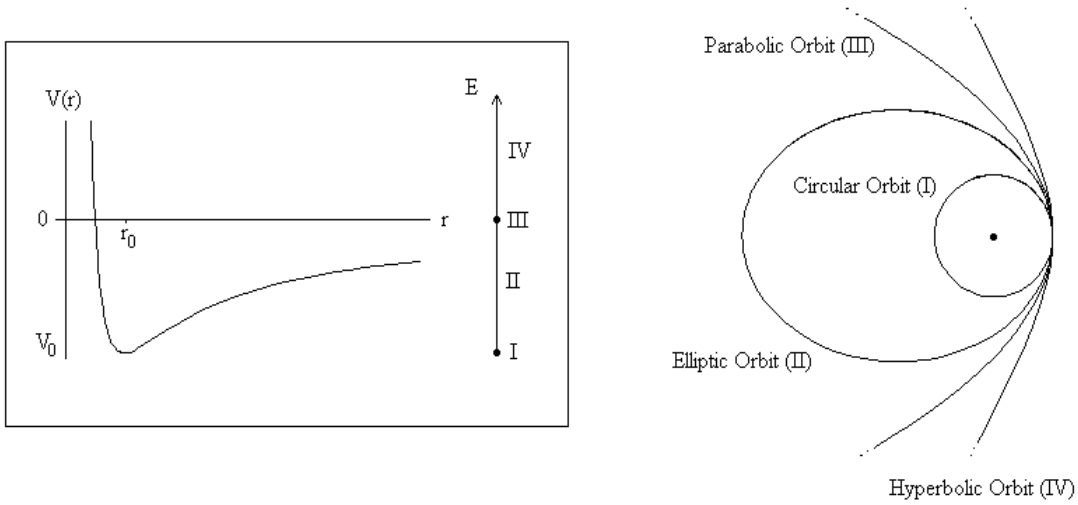


Figure 4.2: Effective potential for the Kepler problem

and we easily verify that $\Delta\theta = 2\pi$, i.e., the bounded orbits of the Kepler problem are closed. This equation can easily be inverted to yield

$$s(\theta) \;=\; s_0\,(1 \;+\; \mathsf{e}\,\cos\theta), \tag{4.18}$$

where we can readily check that this solution also satisfies the radial force equation (4.4).

Kepler's First Law

The solution for $r(\theta)$ is now trivially obtained from $s(\theta)$ as

$$r(\theta) \;=\; \frac{r_0}{1 \;+\; \mathsf{e}\,\cos\theta}, \tag{4.19}$$

where $r_0 = 1/s_0$ denotes the position of the minimum of the *effective* potential $V'(r_0) = 0$. Eq. (4.19) generates an ellipse of semi-major axis $a = r_0/(1-\mathsf{e}^2) = \sqrt{k/(2\,|E|)}$ and semi-minor axis $b = a\,\sqrt{1-\mathsf{e}^2} = \sqrt{\ell^2/(2\mu\,|E|)}$ and, therefore, yields Kepler's First Law.

When we plot the positions of the two objects (of mass m_1 and m_2, respectively) by using Kepler's first law (4.19), with the positions $\mathbf{r}_1$ and $\mathbf{r}_2$ determined by Eqs. (2.27), we obtain the following figures:

Kepler's Second Law

Using Eq. (4.2), we find

$$dt \;=\; \frac{\mu}{\ell}\,r^2\,d\theta \;=\; \frac{2\mu}{\ell}\,dA(\theta),$$

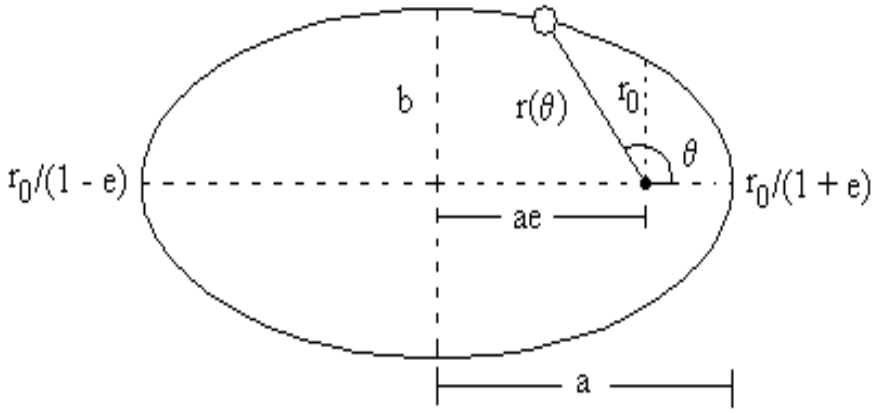


Figure 4.3: Elliptical orbit for the Kepler problem

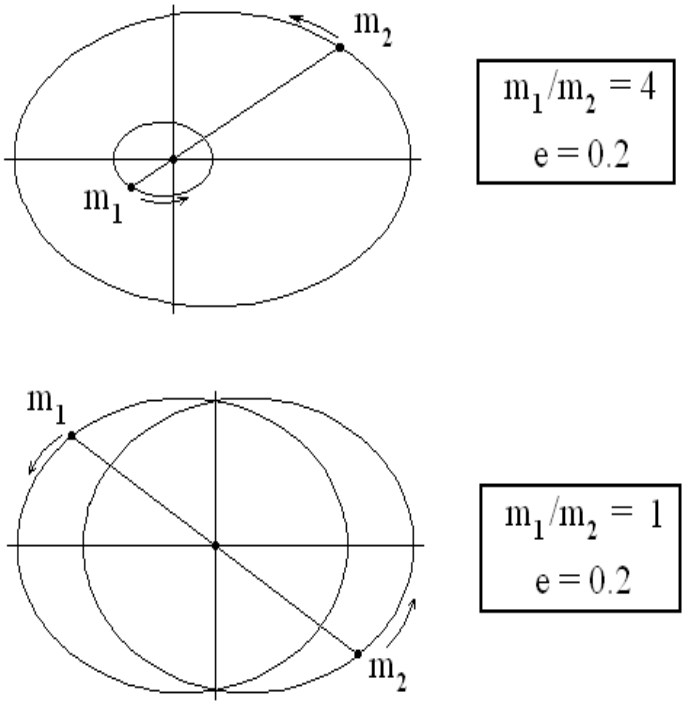


Figure 4.4: Kepler two-body problem

where $dA(\theta) = (\int r\, dr)\, d\theta = \frac{1}{2}\,[r(\theta)]^2\, d\theta$ denotes an infinitesimal area swept by $d\theta$ at radius $r(\theta)$. When integrated, this relation yields Kepler's Second law

$$\Delta t \;=\; \frac{2\mu}{\ell}\;\Delta A, \tag{4.20}$$

i.e., equal areas are swept in equal times since μ and ℓ are constants.

Kepler's Third Law

The orbital period T of a bound system is defined as

$$T \;=\; \int_0^{2\pi}\; \frac{d\theta}{\dot\theta} \;=\; \frac{\mu}{\ell}\;\int_0^{2\pi}\; r^2\, d\theta \;=\; \frac{2\mu}{\ell}\;A \;=\; \frac{2\pi\,\mu}{\ell}\;a\,b$$

where $A = \pi\, ab$ denotes the area of an ellipse with semi-major axis a and semi-minor axis b; here, we used the identity

$$\int_0^{2\pi}\; \frac{d\theta}{(1+\mathsf{e}\,\cos\theta)^2} \;=\; \frac{2\pi}{(1-\mathsf{e}^2)^{3/2}}.$$

Using the expressions for a and b found above, we find

$$T \;=\; \frac{2\pi\,\mu}{\ell}\;\cdot\;\frac{k}{2|E|}\;\cdot\;\sqrt{\frac{\ell^2}{2\mu\,|E|}} \;=\; 2\pi\,\sqrt{\frac{\mu\,k^2}{(2\,|E|)^3}}.$$

If we now substitute the expression for $a = k/2|E|$ and square both sides of this equation, we obtain Kepler's Third Law

$$T^2 \;=\; \frac{(2\pi)^2\mu}{k}\;a^3. \tag{4.21}$$

Note that in Newtonian gravitational theory, where $k/\mu = G\,(m_1 + m_2)$, we find that, although Kepler's Third Law states that T^2/a^3 is a constant for all planets in the solar system, this is only an approximation that holds for $m_1 \gg m_2$ (which holds for all planets).

4.3.2 Unbounded Keplerian Orbits

We now look at the case where the total energy is positive or zero (i.e., $\mathsf{e} \geq 1$). Eq. (4.19) yields $r\,(1+\mathsf{e}\,\cos\theta) = r_0$ or

$$\left(\sqrt{\mathsf{e}^2-1}\;x \;-\; \frac{\mathsf{e}\,r_0}{\sqrt{\mathsf{e}^2-1}}\right)^2 \;-\; y^2 \;=\; \frac{r_0^2}{\mathsf{e}^2-1}.$$

For $\mathsf{e} = 1$, the particle orbit is a parabola $x = (r_0^2 - y^2)/2r_0$, with distance of closest approach at $x(0) = r_0/2$, while for $\mathsf{e} > 1$, the particle orbit is a hyperbola.

4.3.3 Laplace-Runge-Lenz Vector*

Since the orientation of the unperturbed Keplerian ellipse is constant (i.e., it does not precess), it turns out there exists a third constant of the motion for the Kepler problem (in addition to energy and angular momentum); we note, however, that only two of these three invariants are independent.

Let us now investigate this additional constant of the motion for the Kepler problem. First, we consider the time derivative of the vector $\mathbf{p} \times \mathbf{L}$, where the linear momentum $\mathbf{p}$ and angular momentum $\mathbf{L}$ are

$$\mathbf{p} \;=\; \mu \left(\dot{r}\, \widehat{r} \;+\; r\dot{\theta}\, \widehat{\theta} \right) \quad \text{and} \quad \mathbf{L} \;=\; \ell\, \widehat{z} \;=\; \mu r^2 \dot{\theta}\, \widehat{z}.$$

The time derivative of the linear momentum is $\dot{\mathbf{p}} = -\nabla U(r) = -U'(r)\, \widehat{r}$ while the angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ is itself a constant of the motion so that

$$\begin{aligned} \frac{d}{dt}\,(\mathbf{p} \times \mathbf{L}) \;&=\; \frac{d\mathbf{p}}{dt} \times \mathbf{L} \;=\; U'(r)\,(r\,\mathbf{p} \;-\; p_r\,\mathbf{r}) \\ &=\; -\,\mu \dot{\mathbf{r}} \cdot \nabla U\; \mathbf{r} \;+\; \mu \mathbf{r} \cdot \nabla U\; \dot{\mathbf{r}}. \end{aligned}$$

By re-arranging some terms, we find

$$\frac{d}{dt}\,(\mathbf{p} \times \mathbf{L}) \;=\; -\,\frac{d}{dt}\,(\mu\, U\; \mathbf{r}) \;+\; \mu\,(\mathbf{r} \cdot \nabla U \;+\; U)\,\dot{\mathbf{r}},$$

or

$$\frac{d\mathbf{A}}{dt} \;=\; \mu\,(\mathbf{r} \cdot \nabla U \;+\; U)\,\dot{\mathbf{r}}, \tag{4.22}$$

where the vector $\mathbf{A} = \mathbf{p} \times \mathbf{L} + \mu\, U(r)\; \mathbf{r}$ defines the Laplace-Runge-Lenz (LRL) vector. We immediately note that the LRL vector $\mathbf{A}$ is a constant of the motion if the potential $U(r)$ satisfies the condition

$$\mathbf{r} \cdot \nabla U(r) \;+\; U(r) \;=\; \frac{d(r\, U)}{dr} \;=\; 0.$$

For the Kepler problem, with central potential $U(r) \;=\; -\,k/r$, the Laplace-Runge-Lenz (LRL) vector

$$\mathbf{A} \;=\; \mathbf{p} \times \mathbf{L} \;-\; k\mu\, \widehat{r} \;=\; \left(\frac{\ell^2}{r} \;-\; k\mu \right) \widehat{r} \;-\; \ell\, \mu \dot{r}\; \widehat{\theta} \tag{4.23}$$

is, therefore, a constant of the motion since $\mathbf{r} \cdot \nabla U = -\,U$.

Since the vector $\mathbf{A}$ is constant in both magnitude and direction, where

$$|\mathbf{A}|^2 \;=\; 2\mu\, \ell^2 \left(\frac{p^2}{2\mu} \;+\; U \right) \;+\; k^2 \mu^2 \;=\; k^2 \mu^2 \left(1 \;+\; \frac{2\ell^2 E}{\mu k^2} \right) \;=\; k^2 \mu^2 \mathsf{e}^2,$$

we choose its direction to be along the x-axis and its amplitude is determined at the distance of closest approach $r_{min} = r_0/(1 + \mathsf{e})$. We can easily show that $\mathbf{A} \cdot \widehat{r} = |\mathbf{A}|\, \cos\theta$ leads to the Kepler solution

$$r(\theta) \;=\; \frac{r_0}{1 + \mathsf{e}\, \cos\theta},$$

where $r_0 = \ell^2/k\mu$ and the orbit's eccentricity is $\mathsf{e} = |\mathbf{A}|/k\mu$.

Note that if the Keplerian orbital motion is perturbed by the introduction of an additional potential term $\delta U(r)$, we can show that (to lowest order in the perturbation δU)

$$\frac{d\mathbf{A}}{dt} \;=\; (\delta U \;+\; \mathbf{r}\cdot\nabla\delta U)\;\mathbf{p},$$

and, thus using Eq. (4.23)

$$\mathbf{A}\times\frac{d\mathbf{A}}{dt} \;=\; (\delta U \;+\; \mathbf{r}\cdot\nabla\delta U)\;\Big(p^2 \;+\; \mu\,U\Big)\;\mathbf{L},$$

where $U = -\,k/r$ is the unperturbed Kepler potential. Next, using the expression for the unperturbed total energy

$$E \;=\; \frac{p^2}{2\mu} \;+\; U \;=\; -\,\frac{k}{2a},$$

we define the precession frequency

$$\begin{aligned}\omega_{\rm p}(\theta) \;=\; \widehat{\mathsf{z}}\cdot\frac{\mathbf{A}}{|\mathbf{A}|^2}\times\frac{d\mathbf{A}}{dt} \;&=\; (\delta U \;+\; \mathbf{r}\cdot\nabla\delta U)\;\frac{\ell}{(\mu k\mathsf{e})^2}\,\mu\,(2\,E \;-\; U)\\ &=\; (\delta U \;+\; \mathbf{r}\cdot\nabla\delta U)\;\frac{\ell}{(\mu k\mathsf{e})^2}\,\mu k\,\Big(\frac{1}{r} \;-\; \frac{1}{a}\Big).\end{aligned}$$

Hence, using $a = r_0/(1-\mathsf{e}^2)$, the precession frequency is

$$\omega_{\rm p}(\theta) \;=\; \frac{1}{\ell}\;\Big(1 \;+\; \mathsf{e}^{-1}\,\cos\theta\Big)\,(\delta U \;+\; \mathbf{r}\cdot\nabla\delta U)\,.$$

and the net precession shift $\delta\theta$ of the Keplerian orbit over one unperturbed period is

$$\delta\theta \;=\; \int_0^{2\pi}\;\omega_{\rm p}(\theta)\;\frac{d\theta}{\dot\theta} \;=\; \int_0^{2\pi}\;\left(\frac{1+\mathsf{e}^{-1}\,\cos\theta}{1+\mathsf{e}\,\cos\theta}\right)\;\left[\;r\;\frac{d}{dr}\left(\frac{r\;\delta U}{k}\right)\;\right]_{r=r(\theta)}\;d\theta.$$

For example, if $\delta U = -\,\epsilon/r^2$, then $r\;d(r\delta U/k)/dr = \epsilon/kr$ and the net precession shift is

$$\delta\theta \;=\; \frac{\epsilon}{kr_0}\;\int_0^{2\pi}\;\Big(1+\mathsf{e}^{-1}\,\cos\theta\Big)\;d\theta \;=\; 2\pi\;\frac{\epsilon}{kr_0}.$$

Figure 4.5 shows the numerical solution of the perturbed Kepler problem for the case where $\epsilon \simeq kr_0/16$.

4.4 Isotropic Simple Harmonic Oscillator

As a second example of a central potential with closed bounded orbits, we now investigate the case when the central potential is of the form

$$U(r) \;=\; \frac{k}{2}\;r^2 \quad\to\quad \overline{U}(s) \;=\; \frac{\mu k}{2\ell^2\;s^2}. \tag{4.24}$$

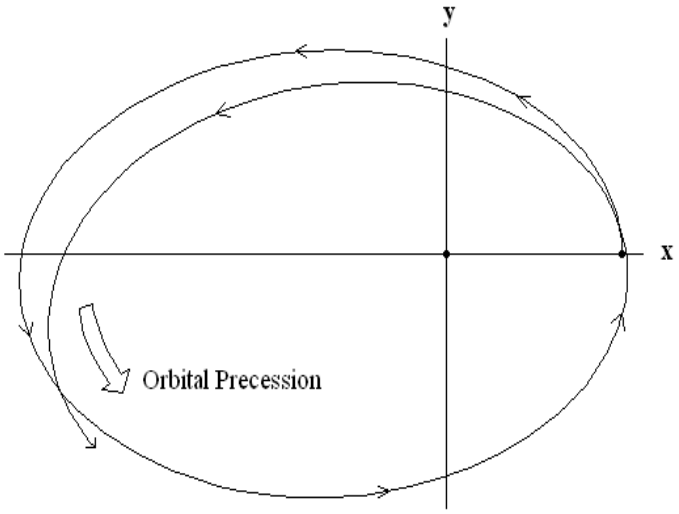


Figure 4.5: Perturbed Kepler problem

The turning points for this problem are expressed as

$$r_1 \;=\; r_0 \left(\frac{1-\mathsf{e}}{1+\mathsf{e}}\right)^{\frac{1}{4}} \;=\; \frac{1}{s_1} \quad \text{and} \quad r_2 \;=\; r_0 \left(\frac{1+\mathsf{e}}{1-\mathsf{e}}\right)^{\frac{1}{4}} \;=\; \frac{1}{s_2},$$

where $r_0 = (\ell^2/\mu k)^{1/4} = 1/s_0$ is the radial position at which the effective potential has a minimum, i.e., $V'(r_0) = 0$ and $V_0 = V(r_0) = k\, r_0^2$ and

$$\mathsf{e} \;=\; \sqrt{1 \;-\; \left(\frac{k r_0^2}{E}\right)^2} \;=\; \sqrt{1 \;-\; \left(\frac{V_0}{E}\right)^2}.$$

Here, we see from Figure 4.6 that orbits are always bounded for $E > V_0$ (and thus $0 \leq \mathsf{e} \leq 1$). Next, using the change of coordinate $q = s^2$ in Eq. (4.10), we obtain

$$\theta \;=\; \frac{-1}{2} \int_{q_2}^{q} \frac{dq}{\sqrt{\varepsilon\, q \;-\; q_0^2 \;-\; q^2}}, \tag{4.25}$$

where $q_2 = (1+\mathsf{e})\,\varepsilon/2$ and $q_0 = s_0^2$. We now substitute $q(\varphi) = (1+\mathsf{e}\,\cos\varphi)\,\varepsilon/2$ in Eq. (4.25) to obtain

$$\theta \;=\; \frac{1}{2}\, \arccos\left[\; \frac{1}{\mathsf{e}} \left(\frac{2\,q}{\varepsilon} - 1\right) \;\right],$$

and we easily verify that $\Delta\theta = \pi$ and bounded orbits are closed. This equation can now be inverted to give

$$r(\theta) \;=\; \frac{r_0\;(1-\mathsf{e}^2)^{1/4}}{\sqrt{1+\mathsf{e}\,\cos 2\theta}}, \tag{4.26}$$

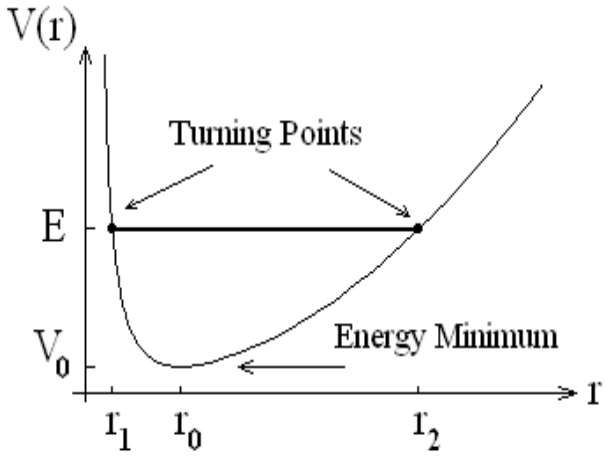


Figure 4.6: Effective potential for the isotropic simple harmonic oscillator problem

which describes the ellipse $x^2/b^2 + y^2/a^2 = 1$, with semi-major axis $a = r_2$ and semi-minor axis $b = r_1$.

The area of the ellipse $A = \pi\, ab = \pi\, r_0^2$ while the *physical* period is

$$T(E,\ell) \;=\; \int_0^{2\pi} \frac{d\theta}{\dot\theta} \;=\; \frac{2\mu\, A}{\ell} \;=\; 2\pi\,\sqrt{\frac{\mu}{k}};$$

note that the *radial* period is $T/2$ since $\Delta\theta = \pi$. We, therefore, find that the period of an isotropic simple harmonic oscillator is independent of the constants of the motion E and ℓ, in analogy with the one-dimensional case.

4.5 Internal Reflection inside a Well

As a last example of bounded motion associated with a central-force potential, we consider the following central potential

$$U(r) \;=\; \begin{cases} -\,U_0 & (r < R) \\ \\ 0 & (r > R) \end{cases}$$

where U_0 is a constant and R denotes the radius of a circle. The effective potential $V(r) = \ell^2/(2\mu r^2) + U(r)$ associated with this potential is shown in Figure 4.7. For energy values

$$V_{\rm min} \;=\; \frac{\ell^2}{2\mu R^2} - U_0 \;<\; E \;<\; V_{\rm max} \;=\; \frac{\ell^2}{2\mu R^2},$$

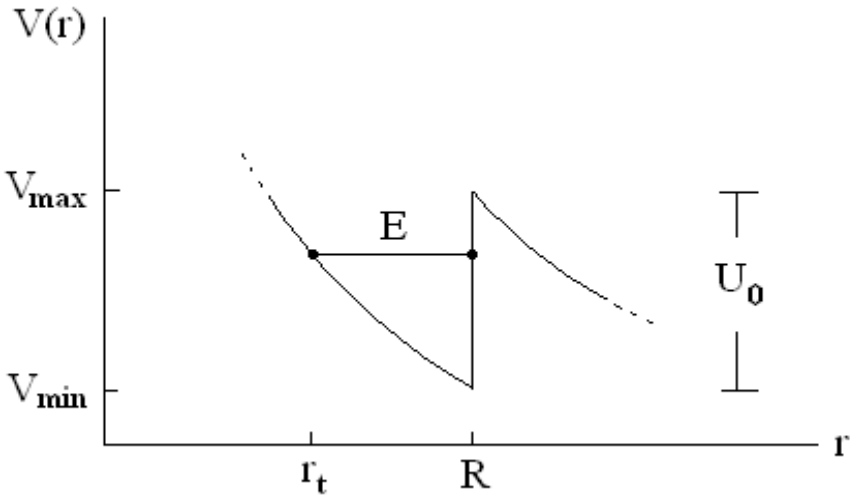


Figure 4.7: Effective potential for the internal hard sphere

Figure 4.7 that bounded motion is possible, with turning points at

$$r_1 \;=\; r_t \;=\; \sqrt{\frac{\ell^2}{2\mu\,(E+U_0)}} \quad \text{and} \quad r_2 \;=\; R.$$

When $E = V_{\min}$, the left turning point reaches its maximum value $r_t = R$ while it reaches its minimum $r_t/R = (1 + U_0/E)^{-\frac{1}{2}} < 1$ when $E = V_{\max}$.

Assuming that the particle starts at $r = r_t$ at $\theta = 0$, the particle orbit is found by integration by quadrature as

$$\theta(s) \;=\; \int_s^{s_t} \frac{d\sigma}{\sqrt{s_t^2 - \sigma^2}},$$

where $s_t = 1/r_t$, which is easily integrated to yield

$$\theta(s) \;=\; \arccos\left(\frac{s}{s_t}\right) \quad \rightarrow \quad r(\theta) \;=\; r_t\,\sec\theta \;\; (\text{for} \;\; \theta \le \Theta),$$

where the maximum angle Θ defines the angle at which the particle hits the turning point R, i.e., $r(\Theta) = R$ and

$$\Theta \;=\; \arccos\left(\sqrt{\frac{\ell^2}{2\mu R^2\,(E+U_0)}}\right).$$

Subsequent motion of the particle involves an infinite sequence of *internal reflections* as shown in Figure 4.8. The case where $E > \ell^2/2\mu R^2$ involves a single turning point and is discussed in Sec. 5.6.2.

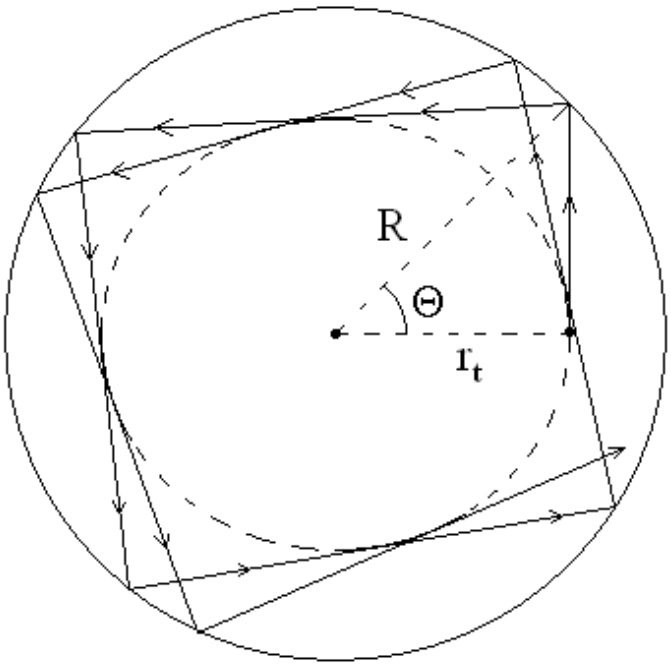
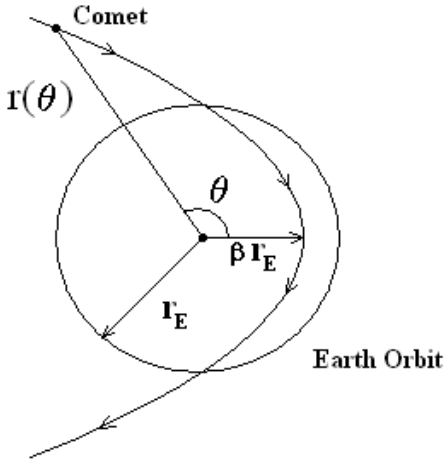


Figure 4.8: Internal reflections inside a hard sphere

4.6 Problems

Problem 1

Consider a comet moving in a parabolic orbit in the plane of the Earth's orbit. If the distance of closest approach of the comet to the sun is $\beta\, r_E$, where r_E is the radius of the Earth's (assumed) circular orbit and where $\beta < 1$, show that the time the comet spends within the orbit of the Earth is given by $\sqrt{2\,(1-\beta)}\,(1+2\,\beta)\;\times\;1\text{ year}/(3\,\pi)$.



Problem 2

Find the force law for a central-force field that allows a particle to move in a spiral orbit given by $r = k\,\theta^2$, where k is a constant.

Problem 3

Consider the perturbed Kepler problem in which a particle of mass m, energy $E < 0$, and angular momentum ℓ is moving in the central-force potential

$$U(r) \;=\; -\,\frac{k}{r} \;+\; \frac{\alpha}{r^2},$$

where the perturbation potential α/r^2 is considered small in the sense that the dimensionless parameter $\epsilon = 2m\alpha/\ell^2 \ll 1$ is small.

(a) Show that the energy equation for this problem can be written using $s = 1/r$ as

$$E \;=\; \frac{\ell^2}{2m}\left[\,(s')^2 \;+\; \gamma^2\,s^2 \;-\; 2\,s_0 s\,\right],$$

where $s_0 = mk/\ell^2$ and $\gamma^2 = 1+\epsilon$.

(b) Show that the turning points are

$$s_1 = \frac{s_0}{\gamma^2}\,(1-\mathsf{e}) \quad \text{and} \quad s_2 = \frac{s_0}{\gamma^2}\,(1+\mathsf{e}),$$

where $\mathsf{e} = \sqrt{1 + 2\,\gamma^2\ell^2 E/mk^2}$.

(c) By solving the integral

$$\theta(s) \;=\; -\int_{s_2}^{s} \frac{d\sigma}{\sqrt{(2mE/\ell^2) + 2\,s_0\sigma - \gamma^2\,\sigma^2}},$$

where $\theta(s_2) = 0$, show that

$$r(\theta) \;=\; \frac{\gamma^2\,r_0}{1 \;+\; \mathsf{e}\,\cos(\gamma\theta)},$$

where $r_0 = 1/s_0$.

Problem 4

A Keplerian elliptical orbit, described by the relation $r(\theta) = r_0/(1+\mathsf{e}\,\cos\theta)$, undergoes a precession motion when perturbed by the perturbation potential $\delta U(r)$, with precession frequency

$$\omega_{\rm p}(\theta) \;=\; \widehat{\mathsf{z}}\cdot\frac{d\mathbf{A}}{dt}\times\frac{\mathbf{A}}{|\mathbf{A}|^2} \;=\; -\,\frac{1}{\ell}\,\Big(1 \;+\; \mathsf{e}^{-1}\,\cos\theta\Big)\,(\delta U \;+\; \mathbf{r}\cdot\nabla\delta U)$$

where $\mathbf{A} = \mathbf{p}\times\mathbf{L} - \mu k\,\widehat{\mathsf{r}}$ denotes the Laplace-Runge-Lenz vector for the unperturbed Kepler problem.

Show that the net precession shift $\delta\theta$ of the Keplerian orbit over one unperturbed period is

$$\delta\theta \;=\; \int_0^{2\pi} \omega_{\rm p}(\theta)\,\frac{d\theta}{\dot\theta} \;=\; -\,6\pi\,\frac{\alpha}{kr_0^2}.$$

if the perturbation potential is $\delta U(r) = -\,\alpha/r^3$, where α is a constant.

Problem 5

Calculate the net precession shift $\delta\theta$ of the Keplerian orbit over one unperturbed period

$$\delta\theta \;=\; \int_0^{2\pi} \omega_{\rm p}(\theta)\,\frac{d\theta}{\dot\theta}$$

if the perturbation potential is $\delta U(r) = -\,\alpha/r^3$.

Chapter 5

Collisions and Scattering Theory

In the previous Chapter, we investigated two types of orbits for two-particle systems evolving under the influence of a central potential. In the present Chapter, we focus our attention on unbounded orbits within the context of *elastic* collision theory. In this context, a collision between two interacting particles involves a three-step process: Step I – two particles are initially infinitely far apart (in which case, the energy of each particle is assumed to be strictly kinetic); Step II – as the two particles approach each other, their interacting potential (repulsive or attractive) causes them to reach a distance of closest approach; and Step III – the two particles then move progressively farther apart (eventually reaching a point where the energy of each particle is once again strictly kinetic).

These three steps form the foundations of Collision *Kinematics* and Collision *Dynamics*. The topic of Collision Kinematics, which describes the collision in terms of the conservation laws of momentum and energy, deals with Steps I and III; here, the incoming particles define the initial state of the two-particle system while the outgoing particles define the final state. The topic of Collision Dynamics, on the other hand, deals with Step II, in which the particular nature of the interaction is taken into account.

5.1 Two-Particle Collisions in the LAB Frame

Consider the collision of two particles (labeled 1 and 2) of masses m_1 and m_2, respectively. Let us denote the velocities of particles 1 and 2 *before* the collision as $\mathbf{u}_1$ and $\mathbf{u}_2$, respectively, while the velocities *after* the collision are denoted $\mathbf{v}_1$ and $\mathbf{v}_2$. Furthermore, the particle momenta before and after the collision are denoted $\mathbf{p}$ and $\mathbf{q}$, respectively.

To simplify the analysis, we define the laboratory (LAB) frame to correspond to the reference frame in which m_2 is at rest (i.e., $\mathbf{u}_2 = 0$); in this collision scenario, m_1 acts as the *projectile* particle and m_2 is the *target* particle. We now write the velocities $\mathbf{u}_1$, $\mathbf{v}_1$, and

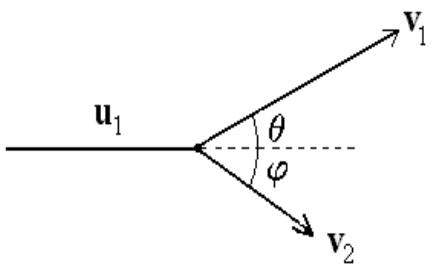


Figure 5.1: Collision kinematics in the LAB frame

$\mathbf{v}_2$ as

$$\left.\begin{array}{rcl} \mathbf{u}_1 &=& u\,\widehat{\mathbf{x}} \\ \mathbf{v}_1 &=& v_1\,(\cos\theta\,\widehat{\mathbf{x}}\,+\,\sin\theta\,\widehat{\mathbf{y}}) \\ \mathbf{v}_2 &=& v_2\,(\cos\varphi\,\widehat{\mathbf{x}}\,-\,\sin\varphi\,\widehat{\mathbf{y}}) \end{array}\right\}, \tag{5.1}$$

where the *deflection* angle θ and the *recoil* angle φ are defined in Figure 5.1. The conservation laws of momentum and energy

$$m_1\mathbf{u}_1 \;=\; m_1\mathbf{v}_1 \,+\, m_2\mathbf{v}_2 \quad\text{and}\quad \frac{m_1}{2}\,u^2 \;=\; \frac{m_1}{2}\,|\mathbf{v}_1|^2 \,+\, \frac{m_2}{2}\,|\mathbf{v}_2|^2$$

can be written in terms of the mass ratio $\alpha = m_1/m_2$ of the projectile mass to the target mass as

$$\alpha\,(u \,-\, v_1\,\cos\theta) \;=\; v_2\,\cos\varphi, \tag{5.2}$$

$$\alpha\,v_1\,\sin\theta \;=\; v_2\,\sin\varphi, \tag{5.3}$$

$$\alpha\,(u^2 \,-\, v_1^2) \;=\; v_2^2. \tag{5.4}$$

Since the **three** equations (5.2)-(5.4) are expressed in terms of **four** unknown quantities $(v_1, \theta,\ v_2, \varphi)$, for given incident velocity u and mass ratio α, we must choose one post-collision coordinate as an independent variable. Here, we choose the recoil angle φ of the target particle, and proceed with finding expressions for $v_1(u, \varphi; \alpha)$, $v_2(u, \varphi; \alpha)$ and $\theta(u, \varphi; \alpha)$.

First, using the square of the mometum components (5.2) and (5.3), we obtain

$$\alpha^2\,v_1^2 \;=\; \alpha^2\,u^2 \;-\; 2\,\alpha\,uv_2\,\cos\varphi \;+\; v_2^2. \tag{5.5}$$

Next, using the energy equation (5.4), we find

$$\alpha^2\,v_1^2 \;=\; \alpha\,\left(\alpha\,u^2 \;-\; v_2^2\right) \;=\; \alpha^2\,u^2 \;-\; \alpha\,v_2^2, \tag{5.6}$$

so that these two equations combine to give

$$v_2(u, \varphi; \alpha) \;=\; 2\left(\frac{\alpha}{1+\alpha}\right)u\,\cos\varphi. \tag{5.7}$$

Once $v_2(u, \varphi; \alpha)$ is known and after substituting Eq. (5.7) into Eq. (5.6), we find

$$v_1(u, \varphi; \alpha) \;=\; u \sqrt{1 \;-\; 4\,\frac{\mu}{M}\,\cos^2\varphi}\,, \tag{5.8}$$

where $\mu/M = \alpha/(1+\alpha)^2$ is the ratio of the reduced mass μ and the total mass M.

Lastly, we take the ratio of the momentum components (5.2) and (5.3) in order to eliminate the unknown v_1 and find

$$\tan\theta \;=\; \frac{v_2\,\sin\varphi}{\alpha\,u \;-\; v_2\,\cos\varphi}.$$

If we substitute Eq. (5.7), we easily obtain

$$\tan\theta \;=\; \frac{2\,\sin\varphi\,\cos\varphi}{1 \;+\; \alpha \;-\; 2\,\cos^2\varphi},$$

or

$$\theta(\varphi; \alpha) \;=\; \arctan\left(\frac{\sin 2\varphi}{\alpha \;-\; \cos 2\varphi}\right). \tag{5.9}$$

In the limit $\alpha = 1$ (i.e., a collision involving identical particles), we find $v_2 = u\cos\varphi$ and $v_1 = u\sin\varphi$ from Eqs. (5.7) and (5.8), respectively, and

$$\tan\theta \;=\; \cot\varphi \quad\rightarrow\quad \varphi \;=\; \frac{\pi}{2} \;-\; \theta,$$

from Eq. (5.9) so that the angular sum $\theta + \varphi$ for like-particle collisions is always 90^o (for $\varphi \neq 0$).

We summarize by stating that after the collision, the momenta $\mathbf{q}_1$ and $\mathbf{q}_2$ in the LAB frame (where m_2 is initially at rest) are

$$\begin{aligned}
\mathbf{q}_1 \;&=\; p\left[1 \;-\; \frac{4\,\alpha}{(1+\alpha)^2}\,\cos^2\varphi\right]^{1/2}\,(\cos\theta\,\hat{\mathbf{x}} \;+\; \sin\theta\,\hat{\mathbf{y}}) \\
\mathbf{q}_2 \;&=\; \frac{2p\,\cos\varphi}{1+\alpha}\,(\cos\varphi\,\hat{\mathbf{x}} \;-\; \sin\varphi\,\hat{\mathbf{y}})
\end{aligned}$$

where $\mathbf{p}_1 = p\,\hat{\mathbf{x}}$ is the initial momentum of particle 1. We note that these expressions for the particle momenta after the collision satisfy the law of conservation of (kinetic) energy in addition to the law of conservation of momentum.

5.2 Two-Particle Collisions in the CM Frame

In the center-of-mass (CM) frame, the elastic collision between particles 1 and 2 is described quite simply; the CM velocities and momenta are, henceforth, denoted with a prime. Before

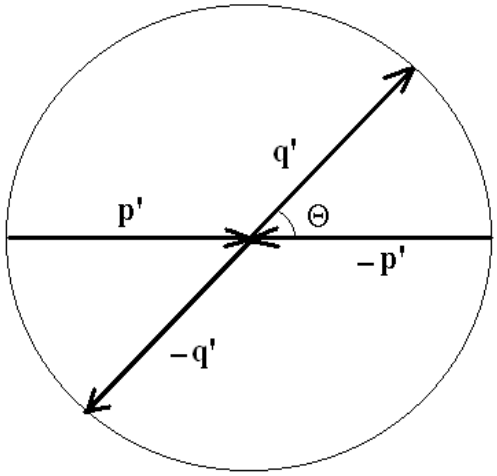


Figure 5.2: Collision kinematics in the CM frame

the collision, the momenta of particles 1 and 2 are equal in magnitude but with opposite directions

$$\mathbf{p}'_1 \;=\; \mu\, u\, \widehat{\mathbf{x}} \;=\; -\,\mathbf{p}'_2,$$

where μ is the reduced mass of the two-particle system. After the collision (see Figure 5.2), conservation of energy-momentum dictates that

$$\mathbf{q}'_1 \;=\; \mu\, u\;(\cos\Theta\,\widehat{\mathbf{x}} \;+\; \sin\Theta\,\widehat{\mathbf{y}}) \;=\; -\,\mathbf{q}'_2,$$

where Θ is the scattering angle in the CM frame and $\mu\, u = p/(1+\alpha)$. Thus the particle velocities after the collision in the CM frame are

$$\mathbf{v}'_1 \;=\; \frac{\mathbf{q}'_1}{m_1} \;=\; \frac{u}{1+\alpha}\;(\cos\Theta\,\widehat{\mathbf{x}} \;+\; \sin\Theta\,\widehat{\mathbf{y}}) \quad \text{and} \quad \mathbf{v}'_2 \;=\; \frac{\mathbf{q}'_2}{m_2} \;=\; -\,\alpha\,\mathbf{v}'_1.$$

It is quite clear, thus, that the initial and final kinematic states lie on the same circle in CM momentum space and the single variable defining the outgoing two-particle state is represented by the CM scattering angle Θ.

5.3 Connection between the CM and LAB Frames

We now establish the connection between the momenta $\mathbf{q}_1$ and $\mathbf{q}_2$ in the LAB frame and the momenta $\mathbf{q}'_1$ and $\mathbf{q}'_2$ in the CM frame. First, we denote the velocity of the CM as

$$\mathbf{w} \;=\; \frac{m_1\,\mathbf{u}_1 + m_2\,\mathbf{u}_2}{m_1+m_2} \;=\; \frac{\alpha\, u}{1+\alpha}\,\widehat{\mathbf{x}},$$

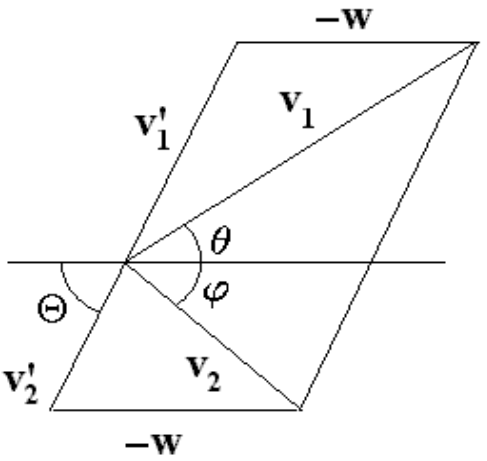


Figure 5.3: CM collision geometry

so that $w = |\mathbf{w}| = \alpha\, u/(1+\alpha)$ and $|\mathbf{v}_2'| = w = \alpha\, |\mathbf{v}_1'|$.

The connection between $\mathbf{v}_1'$ and $\mathbf{v}_1$ is expressed as

$$\mathbf{v}_1' \;=\; \mathbf{v}_1 \;-\; \mathbf{w} \quad\rightarrow\quad \begin{cases} v_1\, \cos\theta &=& w\,(1+\alpha^{-1}\,\cos\Theta) \\ \\ v_1\, \sin\theta &=& w\,\alpha^{-1}\,\sin\Theta \end{cases}$$

so that

$$\tan\theta \;=\; \frac{\sin\Theta}{\alpha+\cos\Theta}, \tag{5.10}$$

and

$$v_1 \;=\; v_1'\,\sqrt{1 \;+\; \alpha^2 \;+\; 2\,\alpha\,\cos\Theta},$$

where $v_1' = u/(1+\alpha)$. Likewise, the connection between $\mathbf{v}_2'$ and $\mathbf{v}_2$ is expressed as

$$\mathbf{v}_2' \;=\; \mathbf{v}_2 \;-\; \mathbf{w} \quad\rightarrow\quad \begin{cases} v_2\, \cos\varphi &=& w\,(1-\cos\Theta) \\ \\ v_2\, \sin\varphi &=& w\,\sin\Theta \end{cases}$$

so that

$$\tan\varphi \;=\; \frac{\sin\Theta}{1-\cos\Theta} \;=\; \cot\frac{\Theta}{2} \quad\rightarrow\quad \varphi \;=\; \frac{1}{2}\,(\pi \;-\; \Theta),$$

and

$$v_2 \;=\; 2\,v_2'\,\sin\frac{\Theta}{2},$$

where $v_2' = \alpha\, u/(1+\alpha) = w$.

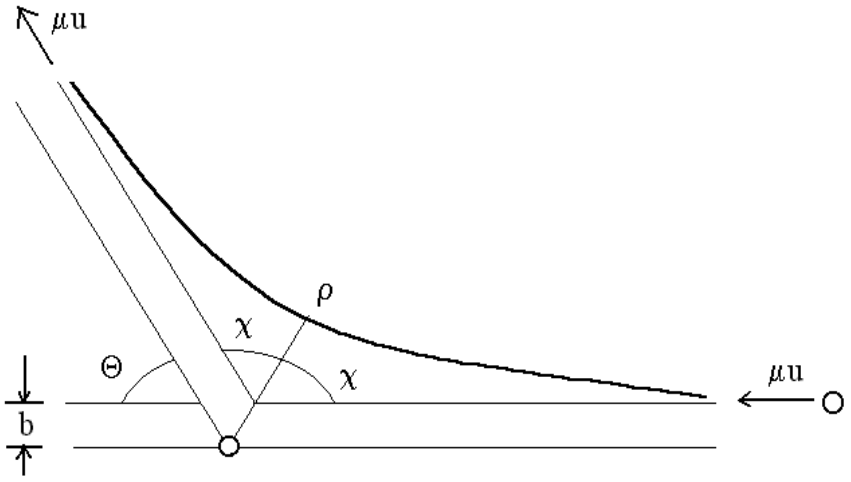


Figure 5.4: Scattering geometry

5.4 Scattering Cross Sections

In the previous Section, we investigated the connection between the initial and final kinematic states of an elastic collision described by Steps I and III, respectively, introduced earlier. In the present Section, we shall investigate Step II, namely, how the distance of closest approach influences the deflection angles (θ, φ) in the LAB frame and Θ in the CM frame.

5.4.1 Definitions

First, we consider for simplicity the case of a projectile particle of mass m being deflected by a repulsive central-force potential $U(r) > 0$ whose center is a rest at the origin (or $\alpha = 0$). As the projectile particle approaches from the right (at $r = \infty$ and $\theta = 0$) moving with speed u, it is progressively deflected until it reaches a minimum radius ρ at $\theta = \chi$ after which the projectile particle moves away from the repulsion center until it reaches $r = \infty$ at a deflection angle $\theta = \Theta$ and again moving with speed u. From Figure 5.4, we can see that the scattering process is symmetric about the line of closest approach (i.e., $2\chi = \pi - \Theta$, where Θ is the CM deflection angle). The angle of closest approach

$$\chi \;=\; \frac{1}{2}\,(\pi - \Theta) \tag{5.11}$$

is a function of the distance of closest approach ρ, the total energy E, and the angular momentum ℓ. The distance ρ is, of course, a turning point ($\dot{r} = 0$) and is the only root of

the equation

$$E \;=\; U(\rho) \;+\; \frac{\ell^2}{2m\,\rho^2}, \tag{5.12}$$

where $E = m\,u^2/2$ is the total initial energy of the projectile particle.

The path of the projectile particle in Figure 5.4 is labeled by the *impact parameter* b (the distance of closest approach in the non-interacting case: $U = 0$) and a simple calculation (using $\mathbf{r} \times \mathbf{v} = bu\,\hat{\mathbf{z}}$) shows that the angular momentum is

$$\ell \;=\; mu\,b \;=\; \sqrt{2m\,E}\;b. \tag{5.13}$$

It is thus quite clear that ρ is a function of E, m, and b. Hence, the angle χ is defined in terms of the standard integral

$$\chi \;=\; \int_\rho^\infty \frac{(\ell/r^2)\,dr}{\sqrt{2m\,[E - U(r)] \;-\; (\ell^2/r^2)}} \;=\; \int_0^{b/\rho} \frac{dx}{\sqrt{1 \;-\; x^2 \;-\; U(b/x)/E}}. \tag{5.14}$$

Once an expression $\Theta(b)$ is obtained from Eq. (5.14), we may invert it to obtain $b(\Theta)$.

5.4.2 Scattering Cross Sections in CM and LAB Frames

We are now ready to discuss the *likelyhood* of the outcome of a collision by introducing the concept of differential cross section $\sigma'(\Theta)$ in the CM frame. The infinitesimal cross section $d\sigma'$ in the CM frame is defined in terms of $b(\Theta)$ as $d\sigma'(\Theta) = \pi\,db^2(\Theta)$. Using Eqs. (5.11) and (5.14), the differential cross section in the CM frame is defined as

$$\sigma'(\Theta) \;=\; \frac{d\sigma'}{2\pi\,\sin\Theta\,d\Theta} \;=\; \frac{b(\Theta)}{\sin\Theta}\,\left|\frac{db(\Theta)}{d\Theta}\right|, \tag{5.15}$$

and the total cross section is, thus, defined as

$$\sigma_T \;=\; 2\pi\,\int_0^\pi \sigma'(\Theta)\,\sin\Theta\,d\Theta.$$

We note that, in Eq. (5.15), the quantity $db/d\Theta$ is often negative and, thus, we must take its absolute value to ensure that $\sigma'(\Theta)$ is positive.

The differential cross section can also be written in the LAB frame in terms of the deflection angle θ as

$$\sigma(\theta) \;=\; \frac{d\sigma}{2\pi\,\sin\theta\,d\theta} \;=\; \frac{b(\theta)}{\sin\theta}\,\left|\frac{db(\theta)}{d\theta}\right|. \tag{5.16}$$

Since the infinitesimal cross section $d\sigma = d\sigma'$ is the same in both frames (i.e., the likelyhood of a collision should not depend on the choice of a frame of reference), we find

$$\sigma(\theta)\,\sin\theta\,d\theta \;=\; \sigma'(\Theta)\,\sin\Theta\,d\Theta,$$

from which we obtain

$$\sigma(\theta) \;=\; \sigma'(\Theta)\,\frac{\sin\Theta}{\sin\theta}\,\frac{d\Theta}{d\theta}, \tag{5.17}$$

or

$$\sigma'(\Theta) \;=\; \sigma(\theta)\,\frac{\sin\theta}{\sin\Theta}\,\frac{d\theta}{d\Theta}. \tag{5.18}$$

Eq. (5.17) yields an expression for the differential cross section in the LAB frame $\sigma(\theta)$ once the differential cross section in the CM frame $\sigma'(\Theta)$ and an explicit formula for $\Theta(\theta)$ are known. Eq. (5.18) represents the inverse transformation $\sigma(\theta) \to \sigma'(\Theta)$. We point out that, whereas the CM differential cross section $\sigma'(\Theta)$ is naturally associated with theoretical calculations, the LAB differential cross section $\sigma(\theta)$ is naturally associated with experimental measurements. Hence, the transformation (5.17) is used to translate a theoretical prediction into an observable experimental cross section, while the transformation (5.18) is used to translate experimental measurements into a format suitable for theoretical analysis.

We note that these transformations rely on finding relations between the LAB deflection angle θ and the CM deflection angle Θ given by Eq. (5.10), which can be converted into

$$\sin(\Theta-\theta) \;=\; \alpha\,\sin\theta. \tag{5.19}$$

For example, using these relations, we now show how to obtain an expression for Eq. (5.17) by using Eqs. (5.10) and (5.19). First, we use Eq. (5.19) to obtain

$$\frac{d\Theta}{d\theta} \;=\; \frac{\alpha\,\cos\theta \;+\; \cos(\Theta-\theta)}{\cos(\Theta-\theta)}, \tag{5.20}$$

where

$$\cos(\Theta-\theta) \;=\; \sqrt{1 \;-\; \alpha^2\,\sin^2\theta}\;.$$

Next, using Eq. (5.10), we show that

$$\begin{aligned}\frac{\sin\Theta}{\sin\theta} \;&=\; \frac{\alpha+\cos\Theta}{\cos\theta} \;=\; \frac{\alpha \;+\; [\cos(\Theta-\theta)\,\cos\theta \;-\; \overbrace{\sin(\Theta-\theta)}^{=\,\alpha\,\sin\theta}\,\sin\theta]}{\cos\theta} \\ &=\; \frac{\alpha\,(1-\sin^2\theta) \;+\; \cos(\Theta-\theta)\,\cos\theta}{\cos\theta} \;=\; \alpha\,\cos\theta \;+\; \sqrt{1 \;-\; \alpha^2\,\sin^2\theta}\;. \end{aligned} \tag{5.21}$$

Thus by combining Eqs. (5.20) and (5.21), we find

$$\frac{\sin\Theta}{\sin\theta}\,\frac{d\Theta}{d\theta} \;=\; \frac{[\alpha\,\cos\theta \;+\; \sqrt{1 \;-\; \alpha^2\,\sin^2\theta}]^2}{\sqrt{1 \;-\; \alpha^2\,\sin^2\theta}} \;=\; 2\,\alpha\,\cos\theta \;+\; \frac{1+\alpha^2\,\cos 2\theta}{\sqrt{1 \;-\; \alpha^2\,\sin^2\theta}}, \tag{5.22}$$

which is valid for $\alpha < 1$. Lastly, noting from Eq. (5.19), that the CM deflection angle is defined as

$$\Theta(\theta) \;=\; \theta \;+\; \arcsin(\,\alpha\,\sin\theta\,),$$

the transformation $\sigma'(\Theta) \to \sigma(\theta)$ is now complete. Similar manipulations yield the transformation $\sigma(\theta) \to \sigma'(\Theta)$. We note that the LAB-frame cross section $\sigma(\theta)$ are generally difficult to obtain for arbitrary mass ratio $\alpha = m_1/m_2$.

5.5 Rutherford Scattering

As an explicit example of the scattering formalism developed in this Chapter, we investigate the scattering of a charged particle of mass m_1 and charge q_1 by another charged particle of mass $m_2 \gg m_1$ and charge q_2 such that $q_1\,q_2 > 0$ and $\mu \simeq m_1$. This situation leads to the two particles experiencing a repulsive central force with potential

$$U(r) \;=\; \frac{k}{r},$$

where $k = q_1 q_2/(4\pi\,\varepsilon_0) > 0$.

The turning-point equation in this case is

$$E \;=\; E\,\frac{b^2}{\rho^2} \;+\; \frac{k}{\rho},$$

whose solution is the distance of closest approach

$$\rho \;=\; r_0 \;+\; \sqrt{r_0^2 + b^2} \;=\; b\left(\epsilon \;+\; \sqrt{1+\epsilon^2}\right), \tag{5.23}$$

where $2\,r_0 = k/E$ is the distance of closest approach for a *head-on* collision (for which the impact parameter b is zero) and $\epsilon = r_0/b$; note, here, that the second solution $r_0 - \sqrt{r_0^2 + b^2}$ to the turning-point equation is negative and, therefore, is not allowed. The problem of the electrostatic repulsive interaction between a positively-charged alpha particle (i.e., the nucleus of a helium atom) and positively-charged nucleus of a gold atom was first studied by Rutherford and the scattering cross section for this problem is known as the Rutherford cross section.

The angle χ at which the distance of closest approach is reached is calculated from Eq. (5.14) as

$$\chi \;=\; \int_0^{b/\rho} \frac{dx}{\sqrt{1 \;-\; x^2 \;-\; 2\,\epsilon\,x}} \;=\; \int_0^{b/\rho} \frac{dx}{\sqrt{(1+\epsilon^2) \;-\; (x+\epsilon)^2}}, \tag{5.24}$$

where

$$\frac{b}{\rho} \;=\; \frac{1}{\epsilon + \sqrt{1+\epsilon^2}} \;=\; -\,\epsilon \;+\; \sqrt{1+\epsilon^2}.$$

Making use of the trigonometric substitution $x = -\,\epsilon + \sqrt{1+\epsilon^2}\,\cos\psi$, we find that

$$\chi \;=\; \arccos\left(\frac{\epsilon}{\sqrt{1+\epsilon^2}}\right) \quad\rightarrow\quad \epsilon \;=\; \cot\chi,$$

which becomes

$$\frac{b}{r_0} \;=\; \tan\chi. \tag{5.25}$$

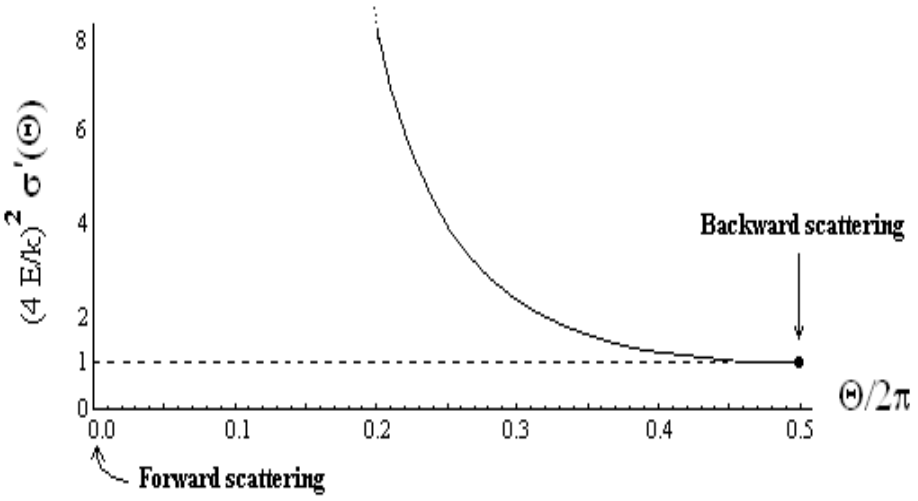


Figure 5.5: Rutherford scattering cross-section

Using the relation (5.11), we now find

$$b(\Theta) \;=\; r_0 \;\cot\frac{\Theta}{2}, \tag{5.26}$$

and thus $db(\Theta)/d\Theta = -\,(r_0/2)\;\csc^2(\Theta/2)$. The CM Rutherford cross section is

$$\sigma'(\Theta) \;=\; \frac{b(\Theta)}{\sin\Theta}\;\left|\frac{db(\Theta)}{d\Theta}\right| \;=\; \frac{r_0^2}{4\;\sin^4(\Theta/2)},$$

or

$$\sigma'(\Theta) \;=\; \left(\frac{k}{4E\;\sin^2(\Theta/2)}\right)^2. \tag{5.27}$$

Note that the Rutherford scattering cross section (5.27) does not depend on the sign of k and is thus valid for both repulsive and attractive interactions. Moreover, we note (see Figure 5.5) that the Rutherford scattering cross section becomes very large in the forward direction $\Theta \rightarrow 0$ (where $\sigma' \rightarrow \Theta^{-4}$) while the differential cross section as $\Theta \rightarrow \pi$ behaves as $\sigma' \rightarrow (k/4E)^2$.

5.6 Hard-Sphere and Soft-Sphere Scattering

Explicit calculations of differential cross sections tend to be very complex for general central potentials and, therefore, prove unsuitable for an undergraduate introductory course in Classical Mechanics. In the present Section, we consider two simple central potentials associated with a uniform central potential $U(r) \neq 0$ confined to a spherical region ($r < R$).

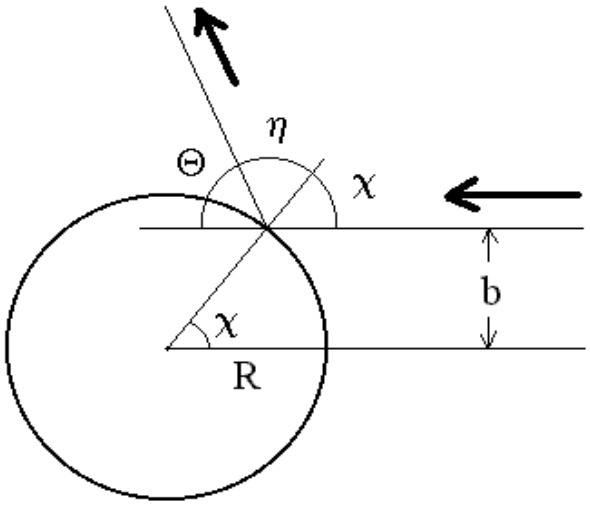


Figure 5.6: Hard-sphere scattering geometry

5.6.1 Hard-Sphere Scattering

We begin by considering the collision of a point-like particle of mass m_1 with a hard sphere of mass m_2 and radius R. In this particular case, the central potential for the hard sphere is

$$U(r) \;=\; \begin{cases} \infty & (\text{for } \; r < R) \\ \\ 0 & (\text{for } \; r > R) \end{cases}$$

and the collision is shown in Figure 5.6. From Figure 5.6, we see that the impact parameter is

$$b \;=\; R\, \sin\chi, \tag{5.28}$$

where χ is the angle of incidence. The angle of reflection η is different from the angle of incidence χ for the case of arbitrary mass ratio $\alpha = m_1/m_2$. To show this, we decompose the velocities in terms of components perpendicular and tangential to the surface of the sphere at the point of impact, i.e., we respectively find

$$\begin{aligned} \alpha\, u\, \cos\chi &= v_2 \;-\; \alpha\, v_1\, \cos\eta \\ \alpha\, u\, \sin\chi &= \alpha\, v_1\, \sin\eta. \end{aligned}$$

From these expressions we obtain

$$\tan\eta \;=\; \frac{\alpha\, u\, \sin\chi}{v_2 - \alpha\, u\, \cos\chi}.$$

From Figure 5.6, we also find the deflection angle $\theta = \pi - (\chi + \eta)$ and the recoil angle $\varphi = \chi$ and thus, according to Chap. 5,

$$v_2 \;=\; \left(\frac{2\,\alpha}{1+\alpha}\right) u\;\cos\chi,$$

and thus

$$\tan\eta \;=\; \left(\frac{1+\alpha}{1-\alpha}\right)\tan\chi. \tag{5.29}$$

We, therefore, easily see that $\eta = \chi$ (the standard form of the Law of Reflection) only if $\alpha = 0$ (i.e., the target particle is infinitely massive).

In the CM frame, the collision is symmetric with a deflection angle $\chi = \frac{1}{2}(\pi - \Theta)$, so that

$$b \;=\; R\,\sin\chi \;=\; R\;\cos\frac{\Theta}{2}.$$

The scattering cross section in the CM frame is

$$\sigma'(\Theta) \;=\; \frac{b(\Theta)}{\sin\Theta}\left|\frac{db(\Theta)}{d\Theta}\right| \;=\; \frac{R\,\cos(\Theta/2)}{\sin\Theta}\;\cdot\;\left|-\frac{R}{2}\;\sin(\Theta/2)\right| \;=\; \frac{R^2}{4}, \tag{5.30}$$

and the total cross section is

$$\sigma_T \;=\; 2\pi\int_0^\pi\;\sigma'(\Theta)\,\sin\Theta\;d\Theta \;=\; \pi\,R^2, \tag{5.31}$$

i.e., the total cross section for the problem of hard-sphere collision is equal to the effective area of the sphere.

The scattering cross section in the LAB frame can also be obtained for the case $\alpha < 1$ using Eqs. (5.17) and (5.22) as

$$\sigma(\theta) \;=\; \frac{R^2}{4}\;\left(2\,\alpha\;\cos\theta\;+\;\frac{1+\alpha^2\;\cos 2\theta}{\sqrt{1\;-\;\alpha^2\;\sin^2\theta}}\right), \tag{5.32}$$

for $\alpha = m_1/m_2 < 1$. The integration of this formula must yield the total cross section

$$\sigma_T \;=\; 2\pi\int_0^\pi\;\sigma(\theta)\,\sin\theta\;d\theta,$$

where $\theta_{max} = \pi$ for $\alpha < 1$.

5.6.2 Soft-Sphere Scattering

We now consider the scattering of a particle subjected to the attractive potential considered in Sec. 4.5

$$U(r) \;=\; \begin{cases} -\,U_0 & (\text{for}\;\; r < R) \\ \\ 0 & (\text{for}\;\; r > R \end{cases} \tag{5.33}$$

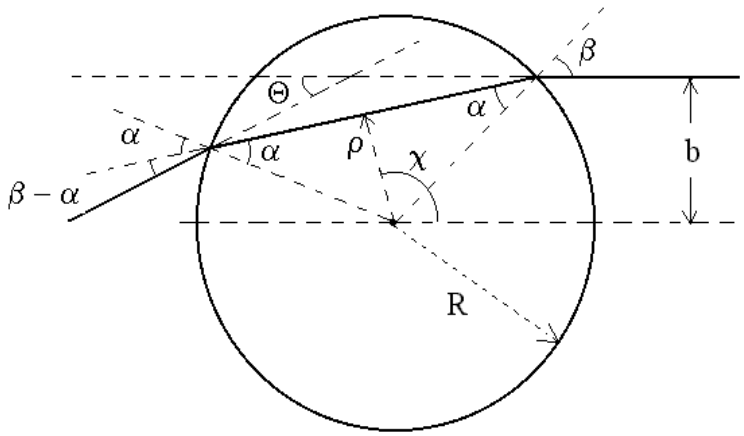


Figure 5.7: Soft-sphere scattering geometry

where the constant U_0 denotes the depth of the attractive potential well and $E > \ell^2/2\mu R^2$ involves a single turning point. We denote β the angle at which the incoming particle enters the *soft-sphere* potential (see Figure 5.7), and thus the impact parameter b of the incoming particle is $b = R\sin\beta$. The particle enters the soft-sphere potential region ($r < R$) and reaches a distance of closest approach ρ, defined from the turning-point condition

$$E \;=\; -\,U_0 \;+\; E\,\frac{b^2}{\rho^2} \quad\rightarrow\quad \rho \;=\; \frac{b}{\sqrt{1+U_0/E}} \;=\; \frac{R}{n}\,\sin\beta,$$

where $n = \sqrt{1+U_0/E}$ denotes the *index of refraction* of the soft-sphere potential region. From Figure 5.7, we note that an optical analogy helps us determine that, through Snell's law, we find

$$\sin\beta \;=\; n\,\sin\left(\beta \;-\; \frac{\Theta}{2}\right), \tag{5.34}$$

where the *transmission* angle α is given in terms of the *incident* angle β and the CM scattering angle $-\,\Theta$ as $\Theta = 2\,(\beta - \alpha)$.

The distance of closest approach is reached at an angle χ is determined as

$$\begin{aligned}\chi \;&=\; \beta \;+\; \int_\rho^R \frac{b\,dr}{r\,\sqrt{n^2r^2-b^2}} \\ &=\; \beta \;+\; \arccos\left(\frac{b}{nR}\right) \;-\; \underbrace{\arccos\left(\frac{b}{n\rho}\right)}_{=\,0}\end{aligned}$$

$$= \beta + \arccos\left(\frac{b}{nR}\right) = \frac{1}{2}(\pi + \Theta), \tag{5.35}$$

and, thus, the impact parameter $b(\Theta)$ can be expressed as

$$b(\Theta) = nR \sin\left(\beta(b) - \frac{\Theta}{2}\right) \quad \rightarrow \quad b(\Theta) = \frac{nR \sin(\Theta/2)}{\sqrt{1 + n^2 - 2n\cos(\Theta/2)}}, \tag{5.36}$$

and its derivative with respect to Θ yields

$$\frac{db}{d\Theta} = \frac{nR}{2} \frac{[n\cos(\Theta/2) - 1]\,[n - \cos(\Theta/2)]}{[1 + n^2 - 2n\cos(\Theta/2)]^{3/2}},$$

and the scattering cross section in the CM frame is

$$\sigma'(\Theta) = \frac{b(\Theta)}{\sin\Theta}\left|\frac{db(\Theta)}{d\Theta}\right| = \frac{n^2R^2}{4} \frac{|[n\cos(\Theta/2) - 1]\,[n - \cos(\Theta/2)]|}{\cos(\Theta/2)\,[1 + n^2 - 2n\cos(\Theta/2)]^2}.$$

Note that, on the one hand, when $\beta = 0$, we find $\chi = \pi/2$ and $\Theta_{\rm min} = 0$, while on the other hand, when $\beta = \pi/2$, we find $b = R$ and

$$1 = n \sin\left(\frac{\pi}{2} - \frac{\Theta_{\rm max}}{2}\right) = n\cos(\Theta_{\rm max}/2) \quad \rightarrow \quad \Theta_{\rm max} = 2\arccos\left(n^{-1}\right).$$

Moreover, when $\Theta = \Theta_{\rm max}$, we find that $db/d\Theta$ vanishes and, therefore, the differential cross section vanishes $\sigma'(\Theta_{\rm max}) = 0$, while at $\Theta = 0$, we find $\sigma'(0) = [n/(n-1)]^2\,(R^2/4)$.

Figure 5.8 shows the soft-sphere scattering cross section $\overline{\sigma}(\Theta)$ (normalized to the hard-sphere cross section $R^2/4$) as a function of Θ for four cases: $n = (1.1, 1.15)$ in the *soft-sphere limit* ($n \rightarrow 1$) and $n = (50, 1000)$ in the *hard-sphere limit* ($n \rightarrow \infty$). We clearly see the strong forward-scattering behavior as $n \rightarrow 1$ (or $U_0 \rightarrow 0$) in the soft-sphere limit and the hard-sphere limit $\overline{\sigma} \rightarrow 1$ as $n \rightarrow \infty$. We note that the total scattering cross section (using the substitution $x = n\cos\Theta/2$)

$$\sigma_T = 2\pi \int_0^{\Theta_{\rm max}} \sigma'(\Theta)\,\sin\Theta\,d\Theta = 2\pi R^2 \int_1^n \frac{(x-1)\,(n^2 - x)\,dx}{(1 + n^2 - 2x)^2} = \pi R^2$$

is independent of the index of refraction n and equals the hard-sphere total cross section (5.31).

The opposite case of a repulsive soft-sphere potential, where $-U_0$ is replaced with U_0 in Eq. (5.33), is treated by replacing $n = (1 + U_0/E)^{\frac{1}{2}}$ with $n = (1 - U_0/E)^{-\frac{1}{2}}$ and Eq. (5.36) is replaced with

$$b(\Theta) = n^{-1} R \sin\left(\beta(b) + \frac{\Theta}{2}\right) \quad \rightarrow \quad b(\Theta) = \frac{R\sin(\Theta/2)}{\sqrt{1 + n^2 - 2n\cos(\Theta/2)}}, \tag{5.37}$$

while Snell's law (5.34) is replaced with

$$\sin\left(\beta + \frac{\Theta}{2}\right) = n \sin\beta.$$

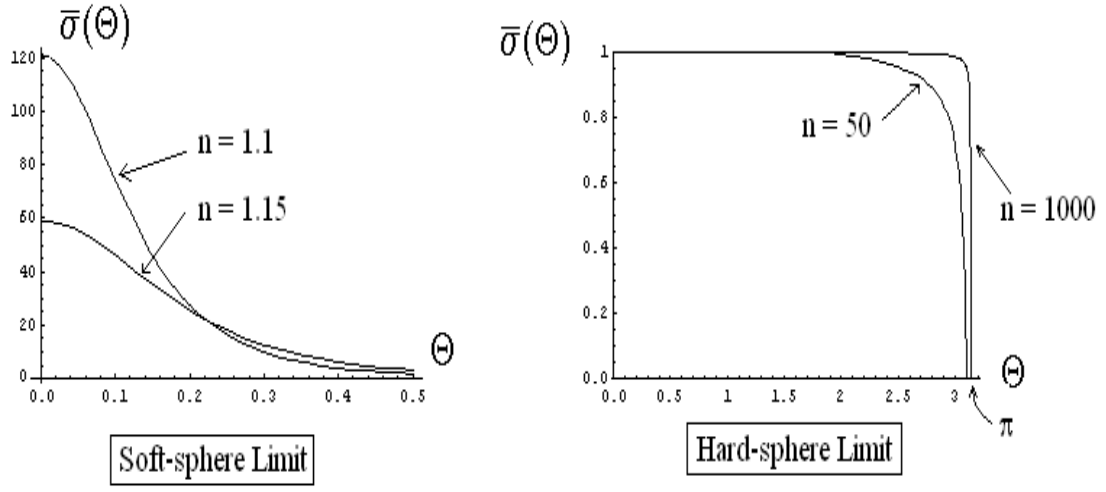


Figure 5.8: Soft-sphere scattering cross-section

5.7 Problems

Problem 1

(a) Using the conservation laws of energy and momentum, solve for $v_1(u, \theta; \beta)$, where $\beta = m_2/m_1$ and

$$\begin{aligned} \mathbf{u}_1 &= u\,\widehat{\mathbf{x}} \\ \mathbf{v}_1 &= v_1\,(\cos\theta\,\widehat{\mathbf{x}} + \sin\theta\,\widehat{\mathbf{y}}) \\ \mathbf{v}_2 &= v_2\,(\cos\varphi\,\widehat{\mathbf{x}} - \sin\varphi\,\widehat{\mathbf{y}}) \end{aligned}$$

(b) Discuss the number of physical solutions for $v_1(u, \theta; \beta)$ for $\beta < 1$ and $\beta > 1$.

(c) For $\beta < 1$, show that physical solutions for $v_1(u, \theta; \beta)$ exist for $\theta < \arcsin(\beta) = \theta_{max}$.

Problem 2

Show that the momentum transfer $\Delta\mathbf{p}_1' = \mathbf{q}_1' - \mathbf{p}_1'$ of the projectile particle in the CM frame has a magnitude

$$|\Delta\mathbf{p}_1'| = 2\,\mu u\,\sin\frac{\Theta}{2},$$

where μ, u, and Θ are the reduced mass, initial projectile LAB speed, and CM scattering

angle, respectively.

Problem 3

Show that the differential cross section $\sigma'(\Theta)$ for the elastic scattering of a particle of mass m from the repulsive central-force potential $U(r) = k/r^2$ with a fixed force-center at $r = 0$ (or an infinitely massive target particle) is

$$\sigma'(\Theta) \;=\; \frac{2\pi^2\,k}{m\,u^2}\;\frac{(\pi-\Theta)}{[\Theta\,(2\pi-\Theta)]^2\;\sin\Theta},$$

where u is the speed of the incoming projectile particle at $r = \infty$.

$$\textit{Hint}: \quad \textit{Show that} \quad b(\Theta) \;=\; \frac{r_0\,(\pi-\Theta)}{\sqrt{2\pi\,\Theta-\Theta^2}}, \quad \textit{where} \quad r_0^2 \;=\; \frac{2\,k}{m\,u^2}.$$

Problem 4

By using the relations $\tan\theta = \sin\Theta/(\alpha+\cos\Theta)$ and/or $\sin(\Theta-\theta) = \alpha\,\sin\theta$, where $\alpha = m_1/m_2$, show that the relation between the differential cross section in the CM frame, $\sigma'(\Theta)$, and the differential cross section in the LAB frame, $\sigma(\theta)$, is

$$\sigma'(\Theta) \;=\; \sigma(\theta)\;\cdot\;\frac{1+\alpha\,\cos\Theta}{(1+2\,\alpha\,\cos\Theta+\alpha^2)^{3/2}}.$$

Problem 5

Consider the scattering of a particle of mass m by the localized attractive central potential

$$U(r) \;=\; \begin{cases} -\,kr^2/2 & r \;\leq\; R \\ \\ 0 & r \;>\; R \end{cases}$$

where the radius R denotes the range of the interaction.

(a) Show that for a particle of energy $E > 0$ moving towards the center of attraction with impact parameter $b = R\,\sin\beta$, the distance of closest approach ρ for this problem is

$$\rho \;=\; \sqrt{\frac{E}{k}\,(\mathsf{e}-1)}, \quad \text{where} \quad \mathsf{e} \;=\; \sqrt{1\;+\;\frac{2\,kb^2}{E}}$$

(b) Show that the angle χ at closest approach is

$$\begin{aligned} \chi &= \beta + \int_\rho^R \frac{(b/r^2)\, dr}{\sqrt{1 - b^2/r^2 + kr^2/2E}} \\ &= \beta + \frac{1}{2} \arccos\left(\frac{2\,\sin^2\beta - 1}{\mathsf{e}}\right) \end{aligned}$$

(c) Using the relation $\chi = \frac{1}{2}(\pi + \Theta)$ between χ and the CM scattering angle Θ, show that

$$\mathsf{e} = \frac{\cos 2\beta}{\cos(2\beta - \Theta)} < 1$$

Chapter 6

Motion in a Non-Inertial Frame

A reference frame is said to be an *inertial* frame if the motion of particles in that frame is subject only to physical forces (i.e., forces are derivable from a physical potential U such that $m\,\ddot{\mathbf{x}} = -\nabla U$). The Principle of Galilean Relativity states that the laws of physics are the same in all inertial frames and that all reference frames moving at constant velocity with respect to an inertial frame are also inertial frames. Hence, physical accelerations are identical in all inertial frames.

In contrast, a reference frame is said to be a *non-inertial* frame if the motion of particles in that frame of reference violates the Principle of Galilean Relativity. Such non-inertial frames include all rotating frames and accelerated reference frames.

6.1 Time Derivatives in Fixed and Rotating Frames

To investigate the relationship between inertial and non-inertial frames, we consider the time derivative of an arbitrary vector $\mathbf{A}$ in two reference frames. The first reference frame is called the *fixed* (inertial) frame and is expressed in terms of the Cartesian coordinates $\mathbf{r}' = (x', y', z')$. The second reference frame is called the *rotating* (non-inertial) frame and is expressed in terms of the Cartesian coordinates $\mathbf{r} = (x, y, z)$. In Figure 6.1, the rotating frame shares the same origin as the fixed frame and the rotation angular velocity $\boldsymbol{\omega}$ of the rotating frame (with respect to the fixed frame) has components $(\omega_x, \omega_y, \omega_z)$.

Since observations can also be made in a rotating frame of reference, we decompose the vector $\mathbf{A}$ in terms of components A_i in the rotating frame (with unit vectors $\hat{\mathsf{x}}^i$). Thus, $\mathbf{A} = A_i\,\hat{\mathsf{x}}^i$ (using the summation rule) and the time derivative of $\mathbf{A}$ as observed in the fixed frame is

$$\frac{d\mathbf{A}}{dt} \;=\; \frac{dA_i}{dt}\,\hat{\mathsf{x}}^i \;+\; A_i\,\frac{d\hat{\mathsf{x}}^i}{dt}. \tag{6.1}$$

The interpretation of the first term is that of the time derivative of $\mathbf{A}$ as observed in the

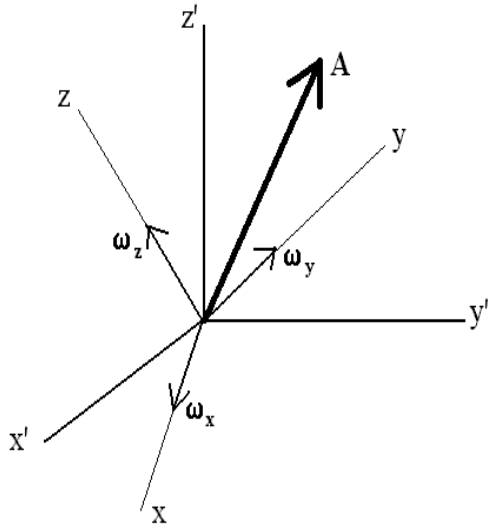


Figure 6.1: Rotating and fixed frames

rotating frame (where the unit vectors $\hat{\mathsf{x}}^i$ are constant) while the second term involves the time-dependence of the relation between the fixed and rotating frames. We now express $d\hat{\mathsf{x}}^i/dt$ as a vector in the rotating frame as

$$\frac{d\hat{\mathsf{x}}^i}{dt} \;=\; R^{ij}\,\hat{\mathsf{x}}_j \;=\; \epsilon^{ijk}\,\omega_k\,\hat{\mathsf{x}}_j, \tag{6.2}$$

where R represents the rotation matrix associated with the rotating frame of reference; this rotation matrix is anti-symmetric ($R^{ij} = -R^{ji}$) and can be written in terms of the anti-symmetric tensor ϵ^{ijk} (defined in terms of the vector product $\mathbf{A}\times\mathbf{B} = A_iB_j\,\epsilon^{ijk}\,\hat{\mathsf{x}}_k$ for two arbitrary vectors $\mathbf{A}$ and $\mathbf{B}$) as $R^{ij} = \epsilon^{ijk}\omega_k$, where ω_k denotes the components of the angular velocity $\boldsymbol{\omega}$ in the rotating frame. Hence, the second term in Eq. (6.1) becomes

$$A_i\,\frac{d\hat{\mathsf{x}}^i}{dt} \;=\; A_i\,\epsilon^{ijk}\,\omega_k\,\hat{\mathsf{x}}_j \;=\; \boldsymbol{\omega}\times\mathbf{A}. \tag{6.3}$$

The time derivative of an arbitrary rotating-frame vector $\mathbf{A}$ in a fixed frame is, therefore, expressed as

$$\left(\frac{d\mathbf{A}}{dt}\right)_f \;=\; \left(\frac{d\mathbf{A}}{dt}\right)_r \;+\; \boldsymbol{\omega}\times\mathbf{A}, \tag{6.4}$$

where $(d/dt)_f$ denotes the time derivative as observed in the fixed (f) frame while $(d/dt)_r$ denotes the time derivative as observed in the rotating (r) frame. An important application of this formula relates to the time derivative of the rotation angular velocity $\boldsymbol{\omega}$ itself. One can easily see that

$$\left(\frac{d\boldsymbol{\omega}}{dt}\right)_f \;=\; \dot{\boldsymbol{\omega}} \;=\; \left(\frac{d\boldsymbol{\omega}}{dt}\right)_r,$$

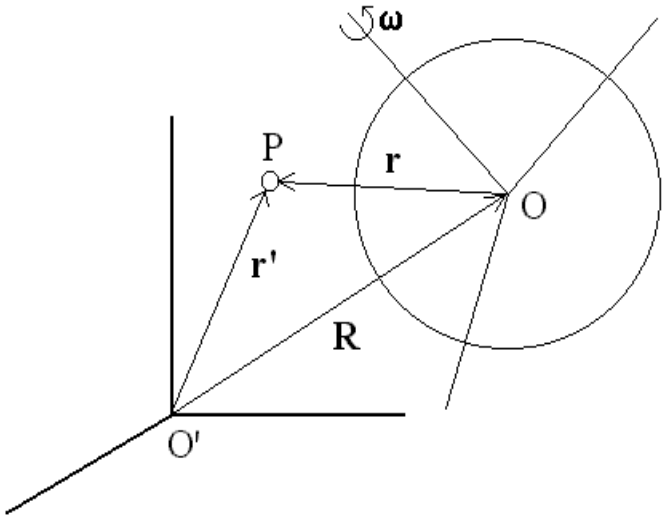


Figure 6.2: General rotating frame

since the second term in Eq. (6.4) vanishes for $\mathbf{A} = \boldsymbol{\omega}$; the time derivative of $\boldsymbol{\omega}$ is, therefore, the same in both frames of reference and is denoted $\dot{\boldsymbol{\omega}}$ in what follows.

6.2 Accelerations in Rotating Frames

We now consider the general case of a rotating frame and fixed frame being related by translation and rotation. In Figure **??**, the position of a point P according to the fixed frame of reference is labeled $\mathbf{r}'$, while the position of the same point according to the rotating frame of reference is labeled $\mathbf{r}$, and

$$\mathbf{r}' \;=\; \mathbf{R} \;+\; \mathbf{r}, \tag{6.5}$$

where $\mathbf{R}$ denotes the position of the origin of the rotating frame according to the fixed frame. Since the velocity of the point P involves the rate of change of position, we must now be careful in defining which time-derivative operator, $(d/dt)_f$ or $(d/dt)_r$, is used.

The velocities of point P as observed in the fixed and rotating frames are defined as

$$\mathbf{v}_f \;=\; \left(\frac{d\mathbf{r}'}{dt}\right)_f \quad \text{and} \quad \mathbf{v}_r \;=\; \left(\frac{d\mathbf{r}}{dt}\right)_r, \tag{6.6}$$

respectively. Using Eq. (6.4), the relation between the fixed-frame and rotating-frame velocities is expressed as

$$\mathbf{v}_f \;=\; \left(\frac{d\mathbf{R}}{dt}\right)_f \;+\; \left(\frac{d\mathbf{r}}{dt}\right)_f \;=\; \mathbf{V} \;+\; \mathbf{v}_r \;+\; \boldsymbol{\omega}\times\mathbf{r}, \tag{6.7}$$

where $\mathbf{V} = (d\mathbf{R}/dt)_f$ denotes the translation velocity of the rotating-frame origin (as observed in the fixed frame).

Using Eq. (6.7), we are now in a position to evaluate expressions for the acceleration of point P as observed in the fixed and rotating frames of reference

$$\mathbf{a}_f \;=\; \left(\frac{d\mathbf{v}_f}{dt}\right)_f \quad \text{and} \quad \mathbf{a}_r \;=\; \left(\frac{d\mathbf{v}_r}{dt}\right)_r, \tag{6.8}$$

respectively. Hence, using Eq. (6.7), we find

$$\begin{aligned}\mathbf{a}_f \;&=\; \left(\frac{d\mathbf{V}}{dt}\right)_f + \left(\frac{d\mathbf{v}_r}{dt}\right)_f + \left(\frac{d\boldsymbol{\omega}}{dt}\right)_f \times \mathbf{r} + \boldsymbol{\omega} \times \left(\frac{d\mathbf{r}}{dt}\right)_f \\ &=\; \mathbf{A} + (\mathbf{a}_r + \boldsymbol{\omega} \times \mathbf{v}_r) + \dot{\boldsymbol{\omega}} \times \mathbf{r} + \boldsymbol{\omega} \times (\mathbf{v}_r + \boldsymbol{\omega} \times \mathbf{r}),\end{aligned}$$

or

$$\mathbf{a}_f \;=\; \mathbf{A} + \mathbf{a}_r + 2\,\boldsymbol{\omega} \times \mathbf{v}_r + \dot{\boldsymbol{\omega}} \times \mathbf{r} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}), \tag{6.9}$$

where $\mathbf{A} = (d\mathbf{V}/dt)_f$ denotes the translational acceleration of the rotating-frame origin (as observed in the fixed frame of reference). We can now write an expression for the acceleration of point P as observed in the rotating frame as

$$\mathbf{a}_r \;=\; \mathbf{a}_f \;-\; \mathbf{A} \;-\; \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) \;-\; 2\,\boldsymbol{\omega} \times \mathbf{v}_r \;-\; \dot{\boldsymbol{\omega}} \times \mathbf{r}, \tag{6.10}$$

which represents the sum of the net *inertial* acceleration $(\mathbf{a}_f - \mathbf{A})$, the centrifugal acceleration $-\,\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})$ and the *Coriolis* acceleration $-\,2\boldsymbol{\omega} \times \mathbf{v}_r$ (see Figures 6.3) and an angular acceleration term $-\,\dot{\boldsymbol{\omega}} \times \mathbf{r}$ which depends explicitly on the time dependence of the rotation angular velocity $\boldsymbol{\omega}$.

The centrifugal acceleration (which is directed outwardly from the rotation axis) represents a familiar *non-inertial* effect in physics. A less familiar *non-inertial* effect is the Coriolis acceleration discovered in 1831 by Gaspard-Gustave Coriolis (1792-1843). Figure 6.3 shows that an object *falling* inwardly also experiences an *eastward* acceleration.

6.3 Lagrangian Formulation of Non-Inertial Motion

We can recover the expression (6.10) for the acceleration in a rotating (non-inertial) frame from a Lagrangian formulation as follows. The Lagrangian for a particle of mass m moving in a non-inertial rotating frame (with its origin coinciding with the fixed-frame origin) in the presence of the potential $U(\mathbf{r})$ is expressed as

$$L(\mathbf{r}, \dot{\mathbf{r}}) \;=\; \frac{m}{2}\,|\dot{\mathbf{r}} + \boldsymbol{\omega} \times \mathbf{r}|^2 \;-\; U(\mathbf{r}), \tag{6.11}$$

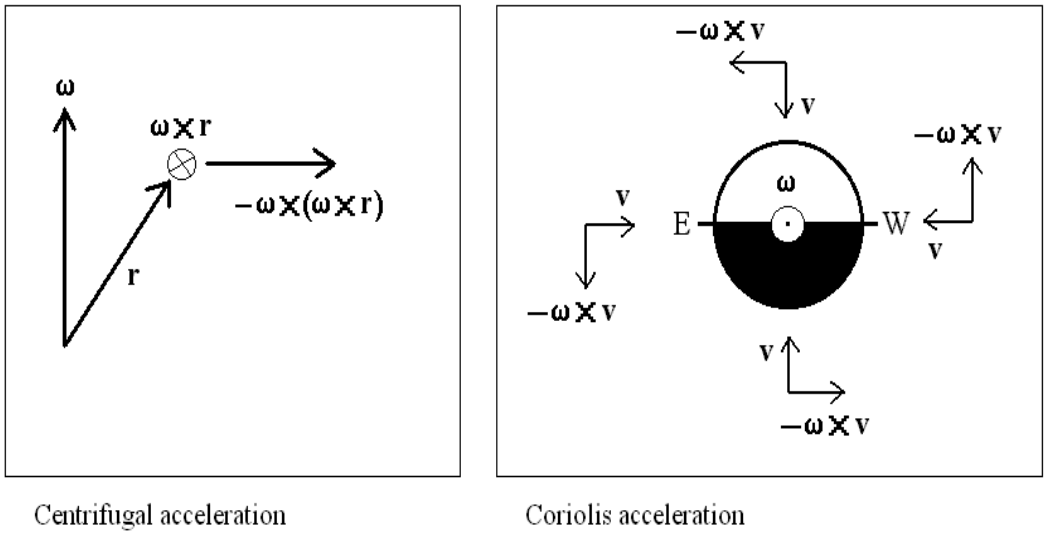


Figure 6.3: Centrifugal and Coriolis accelerations

where $\boldsymbol{\omega}$ is the angular velocity vector and we use the formula

$$|\dot{\mathbf{r}} \,+\, \boldsymbol{\omega} \times \mathbf{r}|^2 \;=\; |\dot{\mathbf{r}}|^2 \,+\, 2\,\boldsymbol{\omega} \cdot (\mathbf{r} \times \dot{\mathbf{r}}) \,+\, \left[\, \omega^2\, r^2 \,-\, (\boldsymbol{\omega} \cdot \mathbf{r})^2 \,\right].$$

Using the Lagrangian (6.11), we now derive the general Euler-Lagrange equation for $\mathbf{r}$. First, we derive an expression for the canonical momentum

$$\mathbf{p} \;=\; \frac{\partial L}{\partial \dot{\mathbf{r}}} \;=\; m\,(\dot{\mathbf{r}} \,+\, \boldsymbol{\omega} \times \mathbf{r})\,, \tag{6.12}$$

and

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{r}}}\right) \;=\; m\,(\ddot{\mathbf{r}} \,+\, \dot{\boldsymbol{\omega}} \times \mathbf{r} \,+\, \boldsymbol{\omega} \times \dot{\mathbf{r}})\,.$$

Next, we derive the partial derivative

$$\frac{\partial L}{\partial \mathbf{r}} \;=\; -\,\nabla U(\mathbf{r}) \;-\; m\,[\,\boldsymbol{\omega} \times \dot{\mathbf{r}} \,+\, \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})\,]\,,$$

so that the Euler-Lagrange equations are

$$m\,\ddot{\mathbf{r}} \;=\; -\,\nabla U(\mathbf{r}) \;-\; m\,[\,\dot{\boldsymbol{\omega}} \times \mathbf{r} \,+\, 2\,\boldsymbol{\omega} \times \dot{\mathbf{r}} \,+\, \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r})\,]\,. \tag{6.13}$$

Here, the potential energy term generates the fixed-frame acceleration, $-\,\nabla U = m\,\mathbf{a}_f$, and thus the Euler-Lagrange equation (6.13) yields Eq. (6.10).

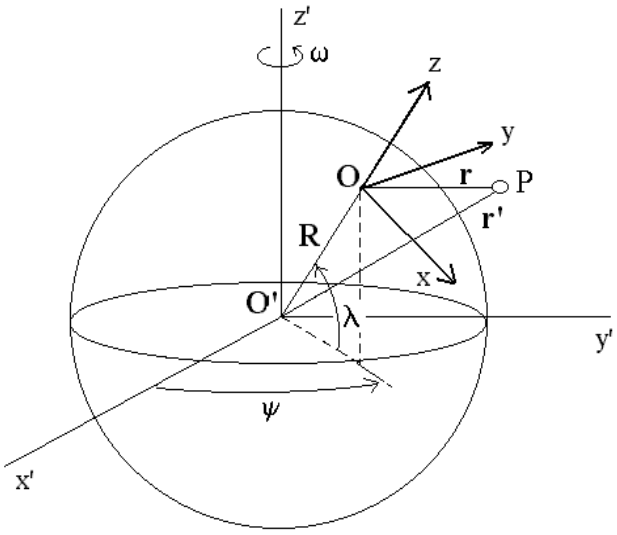


Figure 6.4: Earth frame

6.4 Motion Relative to Earth

We can now apply these non-inertial expressions to the important case of the fixed frame of reference having its origin at the center of Earth (point O' in Figure 6.4) and the rotating frame of reference having its origin at latitude λ and longitude ψ (point O in Figure 6.4). We note that the rotation of the Earth is now represented as $\dot{\psi} = \omega$ and that $\dot{\boldsymbol{\omega}} = 0$.

We arrange the (x, y, z) axis of the rotating frame so that the z-axis is a continuation of the position vector $\mathbf{R}$ of the rotating-frame origin, i.e., $\mathbf{R} = R\,\hat{\mathsf{z}}$ in the rotating frame (where $R = 6378$ km is the radius of a *spherical* Earth). When expressed in terms of the fixed-frame latitude angle λ and the azimuthal angle ψ, the unit vector $\hat{\mathsf{z}}$ is

$$\hat{\mathsf{z}} \;=\; \cos\lambda\;(\cos\psi\,\hat{\mathsf{x}}' \;+\; \sin\psi\,\hat{\mathsf{y}}') \;+\; \sin\lambda\,\hat{\mathsf{z}}',$$

i.e., $\hat{\mathsf{z}}$ points upward. Likewise, we choose the x-axis to be tangent to a *great circle* passing through the North and South poles, so that

$$\hat{\mathsf{x}} \;=\; \sin\lambda\;(\cos\psi\,\hat{\mathsf{x}}' \;+\; \sin\psi\,\hat{\mathsf{y}}') \;-\; \cos\lambda\,\hat{\mathsf{z}}',$$

i.e., $\hat{\mathsf{x}}$ points southward. Lastly, the y-axis is chosen such that

$$\hat{\mathsf{y}} \;=\; \hat{\mathsf{z}} \times \hat{\mathsf{x}} \;=\; -\sin\psi\,\hat{\mathsf{x}}' \;+\; \cos\psi\,\hat{\mathsf{y}}',$$

i.e., $\hat{\mathsf{y}}$ points eastward.

We now consider the acceleration of a point P as observed in the rotating frame O by writing Eq. (6.10) as

$$\frac{d^2\mathbf{r}}{dt^2} \;=\; \mathbf{g}_0 \;-\; \ddot{\mathbf{R}}_f \;-\; \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) \;-\; 2\,\boldsymbol{\omega} \times \frac{d\mathbf{r}}{dt}. \tag{6.14}$$

The first term represents the *pure* gravitational acceleration due to the graviational pull of the Earth on point P (as observed in the fixed frame located at Earth's center)

$$\mathbf{g}_0 \;=\; -\,\frac{G\,M}{|\mathbf{r}'|^3}\,\mathbf{r}',$$

where $\mathbf{r}' = \mathbf{R} + \mathbf{r}$ is the position of point P in the fixed frame and $\mathbf{r}$ is the location of P in the rotating frame. When expressed in terms of rotating-frame spherical coordinates (r, θ, φ):

$$\mathbf{r} \;=\; r\,[\,\sin\theta\;(\cos\varphi\,\hat{\mathsf{x}} \;+\; \sin\varphi\,\hat{\mathsf{y}})\;+\;\cos\theta\,\hat{\mathsf{z}}\,],$$

the fixed-frame position $\mathbf{r}'$ is written as

$$\mathbf{r}' \;=\; (R + r\,\cos\theta)\,\hat{\mathsf{z}} \;+\; r\,\sin\theta\;(\cos\varphi\,\hat{\mathsf{x}} + \sin\varphi\,\hat{\mathsf{y}}),$$

and thus

$$|\mathbf{r}'|^3 \;=\; \left(R^2 \;+\; 2\,R\,r\,\cos\theta \;+\; r^2\right)^{3/2}.$$

The pure gravitational acceleration is, therefore, expressed in the rotating frame of the Earth as

$$\mathbf{g}_0 \;=\; -\,g_0\,\left[\;\frac{(1+\epsilon\,\cos\theta)\,\hat{\mathsf{z}} \;+\; \epsilon\,\sin\theta\,(\cos\varphi\,\hat{\mathsf{x}} + \sin\varphi\,\hat{\mathsf{y}})}{(1 + 2\,\epsilon\,\cos\theta \;+\; \epsilon^2)^{3/2}}\;\right], \tag{6.15}$$

where $g_0 = GM/R^2 = 9.789$ m/s^2 and $\epsilon = r/R \ll 1$.

The angular velocity in the fixed frame is $\boldsymbol{\omega} = \omega\,\hat{\mathsf{z}}'$, where

$$\omega \;=\; \frac{2\pi \text{ rad}}{24 \;\times\; 3600 \text{ sec}} \;=\; 7.27 \;\times\; 10^{-5} \;\text{ rad/s}$$

is the rotation speed of Earth about its axis. In the rotating frame, we find

$$\boldsymbol{\omega} \;=\; \omega\;(\sin\lambda\,\hat{\mathsf{z}} \;-\; \cos\lambda\,\hat{\mathsf{x}}). \tag{6.16}$$

Because the position vector $\mathbf{R}$ rotates with the origin of the rotating frame, its time derivatives yield

$$\begin{aligned}
\dot{\mathbf{R}}_f \;&=\; \boldsymbol{\omega}\times\mathbf{R} \;=\; (\omega\,R\;\cos\lambda)\;\hat{\mathsf{y}},\\
\ddot{\mathbf{R}}_f \;&=\; \boldsymbol{\omega}\times\dot{\mathbf{R}}_f \;=\; \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{R}) \;=\; -\,\omega^2\,R\,\cos\lambda\;(\cos\lambda\,\hat{\mathsf{z}} \;+\; \sin\lambda\,\hat{\mathsf{x}}),
\end{aligned}$$

and thus the centrifugal acceleration due to $\mathbf{R}$ is

$$\ddot{\mathbf{R}}_f \;=\; -\,\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{R}) \;=\; \alpha\,g_0\;\cos\lambda\,(\cos\lambda\,\hat{\mathsf{z}} \;+\; \sin\lambda\,\hat{\mathsf{x}}), \tag{6.17}$$

where $\omega^2 R = 0.0337$ m/s^2 can be expressed in terms of the *pure* gravitational acceleration g_0 as $\omega^2 R \;=\; \alpha\,g_0$, where $\alpha = 3.4 \;\times\; 10^{-3}$. We now define the physical gravitational acceleration as

$$\begin{aligned}
\mathbf{g} \;&=\; \mathbf{g}_0 \;-\; \boldsymbol{\omega}\times[\,\boldsymbol{\omega}\times(\mathbf{R}+\mathbf{r})\,]\\
&=\; g_0\left[\,-\,\left(1 \;-\; \alpha\,\cos^2\lambda\right)\hat{\mathsf{z}} \;+\; (\alpha\;\cos\lambda\,\sin\lambda)\;\hat{\mathsf{x}}\,\right],
\end{aligned} \tag{6.18}$$

where terms of order ϵ have been neglected. For example, a plumb line experiences a small angular deviation $\delta(\lambda)$ (southward) from the true vertical given as

$$\tan\delta(\lambda) \;=\; \frac{g_x}{|g_z|} \;=\; \frac{\alpha\,\sin 2\lambda}{(2-\alpha)+\alpha\,\cos 2\lambda}.$$

This function exhibits a maximum at a latitude $\overline{\lambda}$ defined as $\cos 2\overline{\lambda} = -\alpha/(2-\alpha)$, so that

$$\tan\overline{\delta} \;=\; \frac{\alpha\,\sin 2\overline{\lambda}}{(2-\alpha)+\alpha\,\cos 2\overline{\lambda}} \;=\; \frac{\alpha}{2\sqrt{1-\alpha}} \;\simeq\; 1.7\times 10^{-3},$$

or

$$\overline{\delta} \;\simeq\; 5.86\ \text{arcmin} \quad \text{at} \quad \overline{\lambda} \;\simeq\; \left(\frac{\pi}{4} \;+\; \frac{\alpha}{4}\right)\ \text{rad} \;=\; 45.05^o.$$

We now return to Eq. (6.14), which is written to lowest order in ϵ and α as

$$\frac{d^2\mathbf{r}}{dt^2} \;=\; -\,g\,\widehat{\mathsf{z}} \;-\; 2\,\boldsymbol{\omega}\times\frac{d\mathbf{r}}{dt}, \tag{6.19}$$

where

$$\boldsymbol{\omega}\times\frac{d\mathbf{r}}{dt} \;=\; \omega\,[\,(\dot{x}\,\sin\lambda \;+\; \dot{z}\,\cos\lambda)\,\widehat{\mathsf{y}} \;-\; \dot{y}\,(\sin\lambda\,\widehat{\mathsf{x}} \;+\; \cos\lambda\,\widehat{\mathsf{z}})\,]\,.$$

Thus, we find the three components of Eq. (6.19) written explicitly as

$$\left.\begin{array}{rcl} \ddot{x} &=& 2\,\omega\,\sin\lambda\,\dot{y} \\ \ddot{y} &=& -\,2\,\omega\,(\sin\lambda\,\dot{x} \;+\; \cos\lambda\,\dot{z}) \\ \ddot{z} &=& -\,g \;+\; 2\,\omega\,\cos\lambda\,\dot{y} \end{array}\right\}. \tag{6.20}$$

A first integration of Eq. (6.20) yields

$$\left.\begin{array}{rcl} \dot{x} &=& 2\,\omega\,\sin\lambda\,y \;+\; C_x \\ \dot{y} &=& -\,2\,\omega\,(\sin\lambda\,x \;+\; \cos\lambda\,z) \;+\; C_y \\ \dot{z} &=& -\,g\,t \;+\; 2\,\omega\,\cos\lambda\,y \;+\; C_z \end{array}\right\}, \tag{6.21}$$

where (C_x, C_y, C_z) are constants defined from initial conditions (x_0, y_0, z_0) and $(\dot{x}_0, \dot{y}_0, \dot{z}_0)$:

$$\left.\begin{array}{rcl} C_x &=& \dot{x}_0 \;-\; 2\,\omega\,\sin\lambda\,y_0 \\ C_y &=& \dot{y}_0 \;+\; 2\,\omega\,(\sin\lambda\,x_0 \;+\; \cos\lambda\,z_0) \\ C_z &=& \dot{z}_0 \;-\; 2\,\omega\,\cos\lambda\,y_0 \end{array}\right\}. \tag{6.22}$$

A second integration of Eq. (6.21) yields

$$\begin{aligned} x(t) &= x_0 \;+\; C_x\,t \;+\; 2\,\omega\,\sin\lambda\,\int_0^t y\,dt, \\ y(t) &= y_0 \;+\; C_y\,t \;-\; 2\,\omega\,\sin\lambda\,\int_0^t x\,dt \;-\; 2\,\omega\,\cos\lambda\,\int_0^t z\,dt, \\ z(t) &= z_0 \;+\; C_z\,t \;-\; \frac{1}{2}\,g\,t^2 \;+\; 2\,\omega\,\cos\lambda\,\int_0^t y\,dt, \end{aligned}$$

which can also be rewritten as

$$\left.\begin{array}{rcl} x(t) &=& x_0 \,+\, C_x\, t \,+\, \delta x(t) \\ y(t) &=& y_0 \,+\, C_y\, t \,+\, \delta y(t) \\ z(t) &=& z_0 \,+\, C_z\, t \,-\, \frac{1}{2}\, g\, t^2 \,+\, \delta z(t) \end{array}\right\}, \tag{6.23}$$

where the Coriolis *drifts* are

$$\delta x(t) \;=\; 2\,\omega\,\sin\lambda\left(\, y_0\, t \,+\, \frac{1}{2}\, C_y\, t^2 \,+\, \int_0^t \,\delta y\, dt\,\right) \tag{6.24}$$

$$\begin{aligned} \delta y(t) \;=\;& -\,2\,\omega\,\sin\lambda\,\left(\, x_0\, t \,+\, \frac{1}{2}\, C_x\, t^2 \,+\, \int_0^t \,\delta x\, dt\,\right) \\ & -\,2\,\omega\,\cos\lambda\left(\, z_0\, t \,+\, \frac{1}{2}\, C_z\, t^2 \,-\, \frac{1}{6}\, g\, t^3 \,+\, \int_0^t \,\delta z\, dt\,\right) \end{aligned} \tag{6.25}$$

$$\delta z(t) \;=\; 2\,\omega\,\cos\lambda\left(\, y_0\, t \,+\, \frac{1}{2}\, C_y\, t^2 \,+\, \int_0^t \,\delta y\, dt\,\right). \tag{6.26}$$

Note that each Coriolis drift can be expressed as an infinite series in powers of ω and that all Coriolis effects vanish when $\omega = 0$.

6.4.1 Free-Fall Problem Revisited

As an example of the importance of Coriolis effects in describing motion relative to Earth, we consider the simple *free-fall* problem, where

$$(x_0, y_0, z_0) \;=\; (0, 0, h) \quad \text{and} \quad (\dot{x}_0, \dot{y}_0, \dot{z}_0) \;=\; (0, 0, 0),$$

so that the constants (6.22) are

$$C_x \;=\; 0 \;=\; C_z \quad \text{and} \quad C_y \;=\; 2\,\omega\, h\, \cos\lambda.$$

Substituting these constants into Eqs. (6.23) and keeping only terms up to first order in ω, we find

$$x(t) \;=\; 0, \tag{6.27}$$

$$y(t) \;=\; \frac{1}{3}\, g t^3\, \omega\, \cos\lambda, \tag{6.28}$$

$$z(t) \;=\; h \,-\, \frac{1}{2}\, g t^2. \tag{6.29}$$

Hence, a free-falling object starting from rest touches the ground $z(T) = 0$ after a time $T = \sqrt{2h/g}$ after which time the object has drifted eastward by a distance of

$$y(T) \;=\; \frac{1}{3}\, g T^3\, \omega\, \cos\lambda \;=\; \frac{\omega\, \cos\lambda}{3}\, \sqrt{\frac{8h^3}{g}}.$$

At a height of 100 m and latitude 45^o, we find an eastward drift of 1.55 cm.

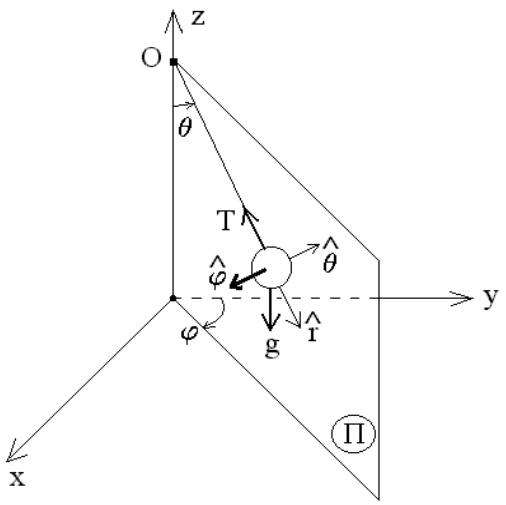


Figure 6.5: Foucault pendulum

6.4.2 Foucault Pendulum

In 1851, Jean Bernard Léon Foucault (1819-1868) was able to demonstrate the role played by Coriolis effects in his investigations of the motion of a pendulum (of length ℓ and mass m) in the rotating frame of the Earth. His analysis showed that, because of the Coriolis acceleration associated with the rotation of the Earth, the motion of the pendulum exhibits a precession motion whose period depends on the latitude at which the pendulum is located.

The equation of motion for the pendulum is given as

$$\ddot{\mathbf{r}} \;=\; \mathbf{a}_f \;-\; 2\,\boldsymbol{\omega}\times\dot{\mathbf{r}}, \tag{6.30}$$

where $\mathbf{a}_f = \mathbf{g}+\mathbf{T}/m$ is the net fixed-frame acceleration of the pendulum expressed in terms of the gravitational acceleration $\mathbf{g}$ and the string tension $\mathbf{T}$ (see Figure 6.5). Note that the vectors $\mathbf{g}$ and $\mathbf{T}$ span a plane Π in which the pendulum moves in the absence of the Coriolis acceleration $-\,2\boldsymbol{\omega}\times\dot{\mathbf{r}}$. Using spherical coordinates (r,θ,φ) in the rotating frame and placing the origin O of the pendulum system at its pivot point (see Figure 6.5), the position of the pendulum bob is

$$\mathbf{r} \;=\; \ell\;[\,\sin\theta\,(\sin\varphi\,\hat{\mathsf{x}} \;+\; \cos\varphi\,\hat{\mathsf{y}}) \;-\; \cos\theta\,\hat{\mathsf{z}}\,] \;=\; \ell\,\hat{r}(\theta,\varphi). \tag{6.31}$$

From this definition, we construct the unit vectors $\widehat{\theta}$ and $\widehat{\varphi}$ as

$$\widehat{\theta} \;=\; \frac{\partial\widehat{r}}{\partial\theta}, \quad \frac{\partial\widehat{r}}{\partial\varphi} \;=\; \sin\theta\,\widehat{\varphi}, \quad \text{and} \quad \frac{\partial\widehat{\theta}}{\partial\varphi} \;=\; \cos\theta\,\widehat{\varphi}. \tag{6.32}$$

Note that, whereas the unit vectors $\hat{r}$ and $\widehat{\theta}$ lie on the plane Π, the unit vector $\widehat{\varphi}$ is perpendicular to it and, thus, the equation of motion of the pendulum *perpendicular* to the

plane Π is

$$\ddot{\mathbf{r}} \cdot \widehat{\varphi} \;=\; -\,2\;(\boldsymbol{\omega} \times \dot{\mathbf{r}}) \cdot \widehat{\varphi}. \tag{6.33}$$

The pendulum velocity is obtained from Eq. (6.31) as

$$\dot{\mathbf{r}} \;=\; \ell \left(\dot{\theta}\, \widehat{\theta} \;+\; \dot{\varphi}\, \sin\theta\, \widehat{\varphi} \right), \tag{6.34}$$

so that the azimuthal component of the Coriolis acceleration is

$$-\,2\;(\boldsymbol{\omega} \times \dot{\mathbf{r}}) \cdot \widehat{\varphi} \;=\; 2\,\ell\omega\,\dot{\theta}\,(\sin\lambda\,\cos\theta \;+\; \cos\lambda\,\sin\theta\,\sin\varphi)\,.$$

If the length ℓ of the pendulum is large, the angular deviation θ of the pendulum can be small enough that $\sin\theta \ll 1$ and $\cos\theta \simeq 1$ and, thus, the azimuthal component of the Coriolis acceleration is approximately

$$-\,2\;(\boldsymbol{\omega} \times \dot{\mathbf{r}}) \cdot \widehat{\varphi} \;\simeq\; 2\,\ell\,(\omega\,\sin\lambda)\;\dot{\theta}. \tag{6.35}$$

Next, the azimuthal component of the pendulum acceleration is

$$\ddot{\mathbf{r}} \cdot \widehat{\varphi} \;=\; \ell \left(\ddot{\varphi}\;\sin\theta \;+\; 2\,\dot{\theta}\,\dot{\varphi}\;\cos\theta \right),$$

which for small angular deviations yields

$$\ddot{\mathbf{r}} \cdot \widehat{\varphi} \;\simeq\; 2\;\ell\;(\dot{\varphi})\;\dot{\theta}. \tag{6.36}$$

By combining these expressions into Eq. (6.33), we obtain an expression for the precession angular frequency of the Foucault pendulum

$$\dot{\varphi} \;=\; \omega\;\sin\lambda \tag{6.37}$$

as a function of latitude λ. As expected, the precession motion is clockwise in the Northern Hemisphere and reaches a maximum at the North Pole ($\lambda = 90^o$). Note that the precession period of the Foucault pendulum is $(1\ \text{day}/\sin\lambda)$ so that the period is 1.41 days at a latitude of 45^o or 2 days at a latitude of 30^o.

The more traditional approach to describing the precession motion of the Foucault pendulum makes use of Cartesian coordinates (x, y, z). The motion of the Foucault pendulum in the (x, y)-plane is described in terms of Eqs. (6.30) as

$$\left.\begin{array}{rcl} \ddot{x} \;+\; \omega_0^2\, x &=& 2\,\omega\sin\lambda\,\dot{y} \\ \ddot{y} \;+\; \omega_0^2\, y &=& -\,2\,\omega\sin\lambda\,\dot{x} \end{array}\right\}, \tag{6.38}$$

where $\omega_0^2 = T/m\ell \simeq g/\ell$ and $\dot{z} \simeq 0$ if ℓ is very large. Figure 6.6 shows the numerical solution of Eqs. (6.38) for the Foucault pendulum starting from rest at $(x_0, y_0) = (0, 1)$ with $2\,(\omega/\omega_0)\,\sin\lambda = 0.05$ at $\lambda = 45^o$. The left figure in Figure 6.6 shows the short time behavior (note the different x and y scales) while the right figure in Figure 6.6 shows the complete Foucault precession. Figure 6.7 shows that, over a finite period of time, the

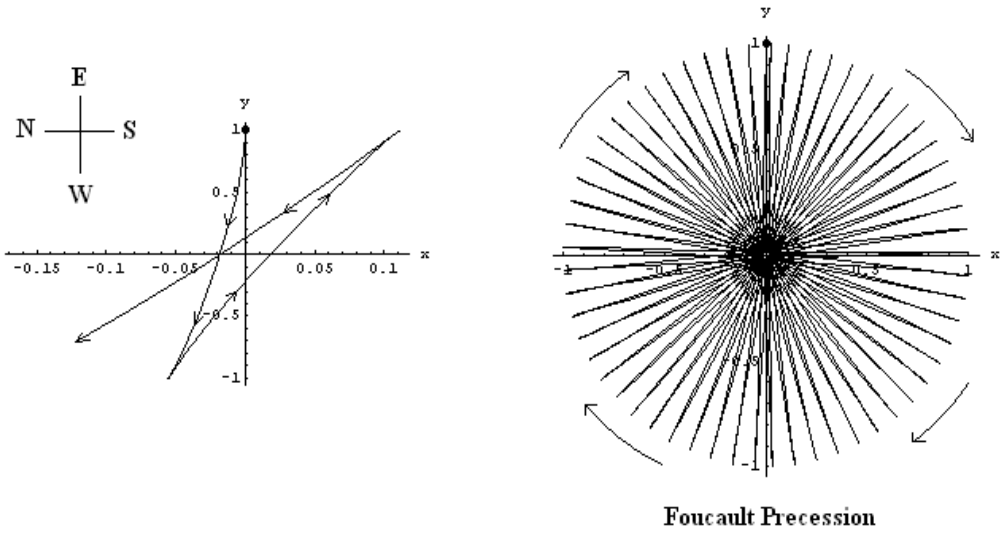


Figure 6.6: Solution of the Foucault pendulum

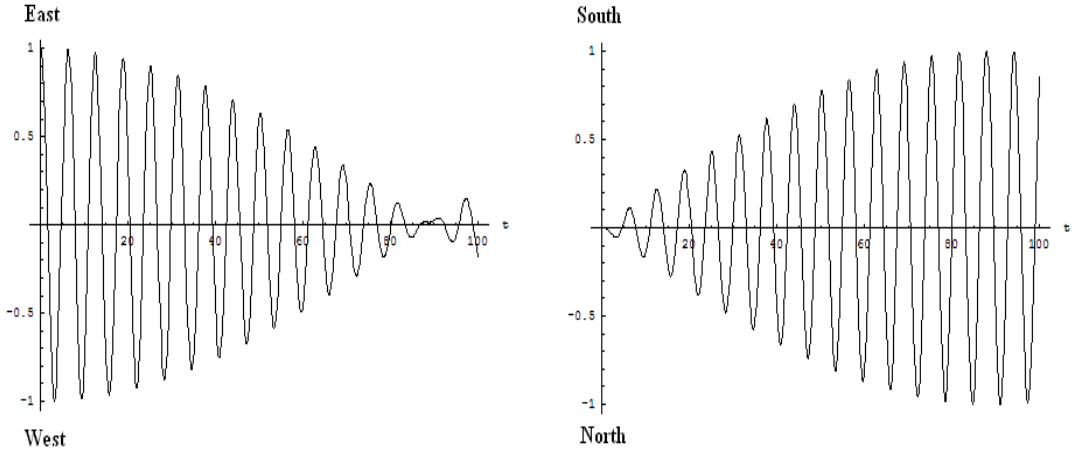


Figure 6.7: Projection of Foucault pendulum

pendulum motion progressively moves from the East-West axis to the North-South axis. We now define the complex-valued function

$$q \;=\; y + i\,x \;=\; \ell\,\sin\theta\,e^{i\varphi}, \tag{6.39}$$

so that Eq. (6.38) becomes

$$\ddot{q} \;+\; \omega_0^2\,q \;-\; 2i\,\omega\sin\lambda\,\dot{q} \;=\; 0.$$

Next, we now insert the eigenfunction $q(t) = \rho\,\exp(i\Omega t)$ into this equation and find that the solution for the eigenfrequency Ω is

$$\Omega \;=\; \omega\,\sin\lambda \;\pm\; \sqrt{\omega^2\,\sin^2\lambda \;+\; \omega_0^2},$$

so that the eigenfunction is

$$q \;=\; \rho\,e^{i\omega\,\sin\lambda\,t}\;\sin\left(\sqrt{\omega^2\,\sin^2\lambda \;+\; \omega_0^2}\;t\right).$$

By comparing this solution with Eq. (6.39), we finally find

$$\rho\;\sin\left(\sqrt{\omega^2\,\sin^2\lambda \;+\; \omega_0^2}\;t\right) \;=\; \ell\;\sin\theta \;\simeq\; \ell\;\theta(t),$$

and

$$\varphi(t) \;=\; (\omega\;\sin\lambda)\;t,$$

from which we recover the Foucault pendulum precession frequency (6.37).

6.5 Problems

Problem 1

(a) Consider the case involving motion on the (x, y)-plane perpendicular to the angular velocity vector $\boldsymbol{\omega} = \omega\,\hat{\mathsf{z}}$ with the potential energy

$$U(\mathbf{r}) \;=\; \frac{1}{2}\,k\,\left(x^2 \;+\; y^2\right).$$

Using the Euler-Lagrange equations (6.13), derive the equations of motion for x and y.

(b) By using the equations of motion derived in Part (a), show that the canonical angular momentum $\ell = \hat{\mathsf{z}}\cdot(\mathbf{r}\times\mathbf{p})$ is a constant of the motion.

Problem 2

If a particle is projected vertically upward to a height h above a point on the Earth's surface at a northern latitude λ, show that it strikes the ground at a point

$$\frac{4\omega}{3}\;\cos\lambda\,\sqrt{\frac{8\,h^3}{g}}$$

to the west. (Neglect air resistance, and consider only small vertical heights.)

Chapter 7

Rigid Body Motion

7.1 Inertia Tensor

7.1.1 Discrete Particle Distribution

We begin our description of rigid body motion by considering the case of a rigid *discrete* particle distribution in which the inter-particle distances are constant. The position of each particle α as measured from a *fixed* laboratory (LAB) frame is

$$\mathbf{r}'_\alpha \;=\; \mathbf{R} \;+\; \mathbf{r}_\alpha,$$

where $\mathbf{R}$ is the position of the center of mass (CM) in the LAB and $\mathbf{r}_\alpha$ is the position of the particle in the CM frame. The velocity of particle α in the LAB frame is

$$\mathbf{v}'_\alpha \;=\; \mathbf{V} \;+\; \boldsymbol{\omega} \times \mathbf{r}_\alpha, \tag{7.1}$$

where $\boldsymbol{\omega}$ is the angular velocity vector associated with the rotation of the particle distribution about an axis of rotation which passes through the CM and $\mathbf{V}$ is the CM velocity in the LAB frame. The total linear momentum in the LAB frame is equal to the momentum of the center of mass since

$$\mathbf{P}' \;=\; \sum_\alpha m_\alpha \mathbf{v}'_\alpha \;=\; M\,\mathbf{V} \;+\; \boldsymbol{\omega} \times \left(\sum_\alpha m_\alpha \mathbf{r}_\alpha \right) \;=\; M\,\mathbf{V},$$

where we have used the definition of the total mass of the particle distribution

$$M \;=\; \sum_\alpha m_\alpha \quad \text{and} \quad \sum_\alpha m_\alpha \mathbf{r}_\alpha \;=\; 0. \tag{7.2}$$

Hence, the total momentum of a rigid body in its CM frame is zero. The total angular momentum in the LAB frame, however, is expressed as

$$\mathbf{L}' \;=\; \sum_\alpha m_\alpha\, \mathbf{r}'_\alpha \times \mathbf{v}'_\alpha \;=\; M\,\mathbf{R} \times \mathbf{V} \;+\; \sum_\alpha m_\alpha\, \mathbf{r}_\alpha \times (\boldsymbol{\omega} \times \mathbf{r}_\alpha)\,, \tag{7.3}$$

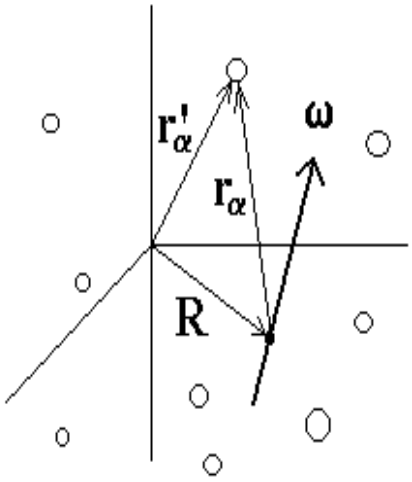


Figure 7.1: Discrete distribution of mass

where we have used the identities (7.2).

The kinetic energy of particle α (with mass m_α) in the LAB frame is

$$K'_\alpha \;=\; \frac{m_\alpha}{2}\,|\mathbf{v}'_\alpha|^2 \;=\; \frac{m_\alpha}{2}\left(|\mathbf{V}|^2 \;+\; 2\,\mathbf{V}\cdot\boldsymbol{\omega}\times\mathbf{r}_\alpha \;+\; |\boldsymbol{\omega}\times\mathbf{r}_\alpha|^2\right),$$

and thus, using the identities (7.2), the total kinetic energy $K' = \sum_\alpha K'_\alpha$ of the particle distribution is

$$K' \;=\; \frac{M}{2}\,|\mathbf{V}|^2 \;+\; \frac{1}{2}\left\{\,\omega^2\left(\sum_\alpha\, m_\alpha\, r_\alpha^2\right) \;-\; \boldsymbol{\omega}\boldsymbol{\omega} \;:\; \left(\sum_\alpha\, m_\alpha\,\mathbf{r}_\alpha\mathbf{r}_\alpha\right)\right\}. \tag{7.4}$$

Looking at Eqs. (7.3) and (7.4), we now introduce the *inertia tensor* of the particle distribution

$$\mathbf{I} \;=\; \sum_\alpha\, m_\alpha\left(r_\alpha^2\,\underline{\mathbf{1}} \;-\; \mathbf{r}_\alpha\,\mathbf{r}_\alpha\right), \tag{7.5}$$

where $\underline{\mathbf{1}}$ denotes the unit tensor (i.e., in Cartesian coordinates, $\underline{\mathbf{1}} = \hat{\mathsf{x}}\hat{\mathsf{x}} + \hat{\mathsf{y}}\hat{\mathsf{y}} + \hat{\mathsf{z}}\hat{\mathsf{z}}$). The inertia tensor can also be represented as a matrix

$$\mathbf{I} \;=\; \begin{pmatrix} \sum_\alpha\, m_\alpha\,(y_\alpha^2+z_\alpha^2) & -\sum_\alpha\, m_\alpha\,(x_\alpha y_\alpha) & -\sum_\alpha\, m_\alpha\,(x_\alpha z_\alpha) \\ -\sum_\alpha\, m_\alpha\,(y_\alpha x_\alpha) & \sum_\alpha\, m_\alpha\,(x_\alpha^2+z_\alpha^2) & -\sum_\alpha\, m_\alpha\,(y_\alpha z_\alpha) \\ -\sum_\alpha\, m_\alpha\,(z_\alpha x_\alpha) & -\sum_\alpha\, m_\alpha\,(z_\alpha y_\alpha) & \sum_\alpha\, m_\alpha\,(x_\alpha^2+y_\alpha^2) \end{pmatrix}, \tag{7.6}$$

where the symmetry property of the inertia tensor ($I^{ji} = I^{ij}$) is readily apparent. In terms of the inertia tensor (7.5), the angular momentum of a rigid body in the CM frame and its

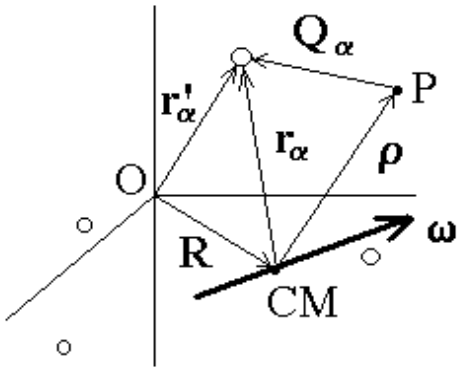


Figure 7.2: Parallel-axes theorem

rotational kinetic energy are

$$\mathbf{L} \;=\; \mathbf{I}\cdot\boldsymbol{\omega} \quad \text{and} \quad K_{rot} \;=\; \frac{1}{2}\,\boldsymbol{\omega}\cdot\mathbf{I}\cdot\boldsymbol{\omega}. \tag{7.7}$$

7.1.2 Parallel-Axes Theorem

A translation of the origin from which the inertia tensor is calculated leads to a different inertia tensor. Let $\mathbf{Q}_\alpha$ denote the position of particle α in a new frame of reference (with its origin located at point P in Figure 7.2) and let $\boldsymbol{\rho} = \mathbf{r}_\alpha - \mathbf{Q}_\alpha$ is the displacement from point CM to point P. The new inertia tensor

$$\mathbf{J} \;=\; \sum_\alpha m_\alpha \left(Q_\alpha^2\,\underline{\mathbf{1}} \;-\; \mathbf{Q}_\alpha\,\mathbf{Q}_\alpha\right)$$

can be expressed as

$$\begin{aligned}\mathbf{J} \;=\; & \sum_\alpha m_\alpha \left(\rho^2\,\underline{\mathbf{1}} \;-\; \boldsymbol{\rho}\,\boldsymbol{\rho}\right) \;+\; \sum_\alpha m_\alpha \left(r_\alpha^2\,\underline{\mathbf{1}} \;-\; \mathbf{r}_\alpha\,\mathbf{r}_\alpha\right) \\ & -\; \left\{\boldsymbol{\rho}\cdot\left(\sum_\alpha m_\alpha\,\mathbf{r}_\alpha\right)\right\}\underline{\mathbf{1}} \;+\; \left\{\boldsymbol{\rho}\,\left(\sum_\alpha m_\alpha\,\mathbf{r}_\alpha\right) \;+\; \left(\sum_\alpha m_\alpha\,\mathbf{r}_\alpha\right)\,\boldsymbol{\rho}\right\}.\end{aligned}$$

Since $M = \sum_\alpha m_\alpha$ and $\sum_\alpha m_\alpha \mathbf{r}_\alpha = 0$, we find

$$\mathbf{J} \;=\; M\,\left(\rho^2\,\underline{\mathbf{1}} \;-\; \boldsymbol{\rho}\,\boldsymbol{\rho}\right) \;+\; \mathbf{I}_{CM}. \tag{7.8}$$

Hence, once the inertia tensor $\mathbf{I}_{CM}$ is calculated in the CM frame, it can be calculated anywhere else. Eq. (7.8) is known as the Parallel-Axes Theorem.

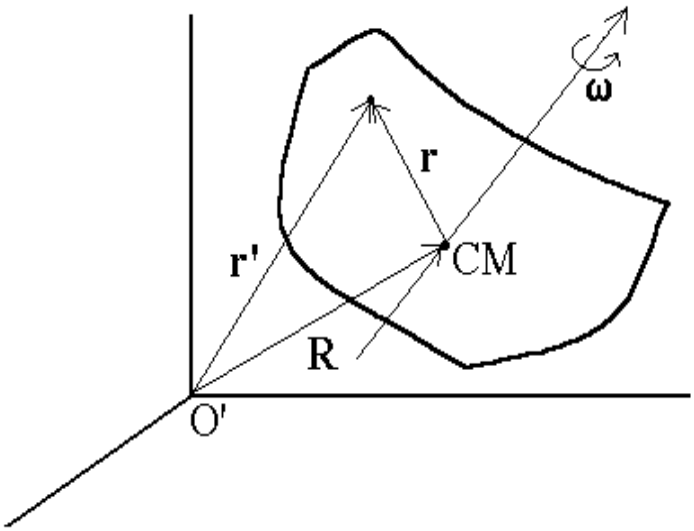


Figure 7.3: Continuous distribution of mass

7.1.3 Continuous Particle Distribution

For a continuous particle distribution the inertia tensor (7.5) becomes

$$\mathbf{I} \;=\; \int \, dm \, \left(r^2 \, \underline{\mathbf{1}} \;-\; \mathbf{r}\mathbf{r} \right), \tag{7.9}$$

where $dm(\mathbf{r}) \;=\; \rho(\mathbf{r})\, d^3r$ is the infinitesimal mass element at point $\mathbf{r}$, with mass density $\rho(\mathbf{r})$.

Consider, for example, the case of a uniform cube of mass M and volume b^3, with $dm = (M/b^3)\, dx\, dy\, dz$. The inertia tensor (7.9) in the LAB frame (with the origin placed at one of its corners) has the components

$$J^{11} \;=\; \frac{M}{b^3} \int_0^b dx \int_0^b dy \int_0^b dz \;\cdot\; \left(y^2 + z^2 \right) \;=\; \frac{2}{3}\, M\, b^2 \;=\; J^{22} \;=\; J^{33} \tag{7.10}$$

$$J^{12} \;=\; -\, \frac{M}{b^3} \int_0^b dx \int_0^b dy \int_0^b dz \;\cdot\; x\, y \;=\; -\, \frac{1}{4}\, M\, b^2 \;=\; J^{23} \;=\; J^{31} \tag{7.11}$$

and thus the inertia matrix for the cube is

$$\mathbf{J} \;=\; \frac{M\, b^2}{12} \begin{pmatrix} 8 & -3 & -3 \\ -3 & 8 & -3 \\ -3 & -3 & 8 \end{pmatrix}. \tag{7.12}$$

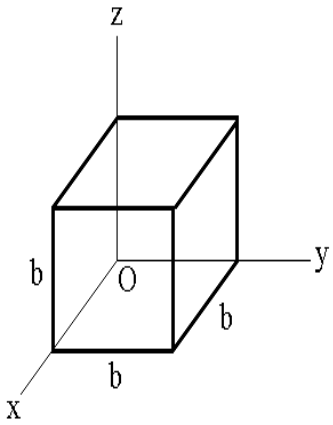


Figure 7.4: Cube

On the other hand, the inertia tensor calculated in the CM frame has the components

$$I^{11} \;=\; \frac{M}{b^3}\int_{-\,b/2}^{b/2} dx \int_{-\,b/2}^{b/2} dy \int_{-\,b/2}^{b/2} dz \;\cdot\; \left(y^2+z^2\right) \;=\; \frac{1}{6}\,M\,b^2 \;=\; I^{22} \;=\; I^{33} \tag{7.13}$$

$$I^{12} \;=\; -\,\frac{M}{b^3}\int_{-\,b/2}^{b/2} dx \int_{-\,b/2}^{b/2} dy \int_{-\,b/2}^{b/2} dz \;\cdot\; x\,y \;=\; 0 \;=\; I^{23} \;=\; I^{31} \tag{7.14}$$

and thus the CM inertia matrix for the cube is

$$\mathbf{I} \;=\; \frac{M\,b^2}{6}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{7.15}$$

The displacement vector $\boldsymbol{\rho}$ from the CM point to the corner O is given as

$$\boldsymbol{\rho} = -\,\frac{b}{2}\,\left(\hat{\mathsf{x}}+\hat{\mathsf{y}}+\hat{\mathsf{z}}\right),$$

so that $\rho^2 = 3b^2/4$. By using the Parallel-Axis Theorem (7.8), the inertia tensor

$$M\,\left(\rho^2\,\underline{\mathbf{1}} \;-\; \boldsymbol{\rho}\,\boldsymbol{\rho}\right) \;=\; \frac{M\,b^2}{4}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

when added to the CM inertia tensor (7.15), yields the inertia tensor (7.12)

7.1.4 Principal Axes of Inertia

In general, the CM inertia tensor $\mathbf{I}$ can be made into a *diagonal* tensor with components given by the eigenvalues I_1, I_2, and I_3 of the inertia tensor. These components (known as principal moments of inertia) are the three roots of the cubic polynomial

$$I^3 \;-\; \mathrm{Tr}(\mathbf{I})\, I^2 \;+\; \mathrm{Ad}(\mathbf{I})\, I \;-\; \mathrm{Det}(\mathbf{I}) \;=\; 0, \tag{7.16}$$

obtained from $\mathrm{Det}(\mathbf{I} - I\,\underline{\mathbf{1}}) = 0$, with coefficients

$$\begin{aligned}
\mathrm{Tr}(\mathbf{I}) &= I^{11} \;+\; I^{22} \;+\; I^{33}, \\
\mathrm{Ad}(\mathbf{I}) &= \mathrm{ad}_{11} \;+\; \mathrm{ad}_{22} \;+\; \mathrm{ad}_{33}, \\
\mathrm{Det}(\mathbf{I}) &= I^{11}\,\mathrm{ad}_{11} \;-\; I^{12}\,\mathrm{ad}_{12} \;+\; I^{13}\,\mathrm{ad}_{13},
\end{aligned}$$

where ad_{ij} is the determinant of the two-by-two matrix obtained from $\mathbf{I}$ by removing the i^{th}-row and j^{th}-column from the inertia matrix $\mathbf{I}$.

Each principal moment of inertia I_i represents the moment of inertia calculated about the principal axis of inertia with unit vector $\hat{\mathsf{e}}_i$. The unit vectors $(\hat{\mathsf{e}}_1, \hat{\mathsf{e}}_2, \hat{\mathsf{e}}_3)$ form a new frame of reference known as the *Body* frame. The unit vectors $(\hat{\mathsf{e}}_1, \hat{\mathsf{e}}_2, \hat{\mathsf{e}}_3)$ are related by a sequence of rotations to the Cartesian CM unit vectors $(\hat{\mathsf{x}}^1, \hat{\mathsf{x}}^2, \hat{\mathsf{x}}^3)$ by the relation

$$\hat{\mathsf{e}}_i \;=\; R_{ij}\,\hat{\mathsf{x}}^j, \tag{7.17}$$

where R_{ij} are components of the rotation matrix R. By denoting as $\mathbf{I}'$ the diagonal inertia tensor calculated in the *body* frame of reference (along the principal axes), we find

$$\mathbf{I}' \;=\; \mathsf{R}\cdot\mathbf{I}\cdot\mathsf{R}^{\mathrm{T}} \;=\; \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix}, \tag{7.18}$$

where R^{T} denotes the transpose of R, i.e., $(\mathsf{R}^{\mathrm{T}})_{ij} = R_{ji}$. In the body frame, the inertia tensor is, therefore, expressed in dyadic form as

$$\mathbf{I}' \;=\; I_1\,\hat{\mathsf{e}}_1\,\hat{\mathsf{e}}_1 \;+\; I_2\,\hat{\mathsf{e}}_2\,\hat{\mathsf{e}}_2 \;+\; I_3\,\hat{\mathsf{e}}_3\,\hat{\mathsf{e}}_3, \tag{7.19}$$

and the rotational kinetic energy (7.7) is

$$K'_{rot} \;=\; \frac{1}{2}\,\boldsymbol{\omega}\cdot\mathbf{I}'\cdot\boldsymbol{\omega} \;=\; \frac{1}{2}\left(I_1\,\omega_1^2 \;+\; I_2\,\omega_2^2 \;+\; I_3\,\omega_3^2\right). \tag{7.20}$$

Note that general rotation matrices have the form

$$\mathsf{R}_n(\alpha) \;=\; \hat{\mathsf{n}}\,\hat{\mathsf{n}} \;+\; \cos\alpha\,(\underline{\mathbf{1}} \;-\; \hat{\mathsf{n}}\,\hat{\mathsf{n}}) \;-\; \sin\alpha\,\hat{\mathsf{n}}\times\underline{\mathbf{1}}, \tag{7.21}$$

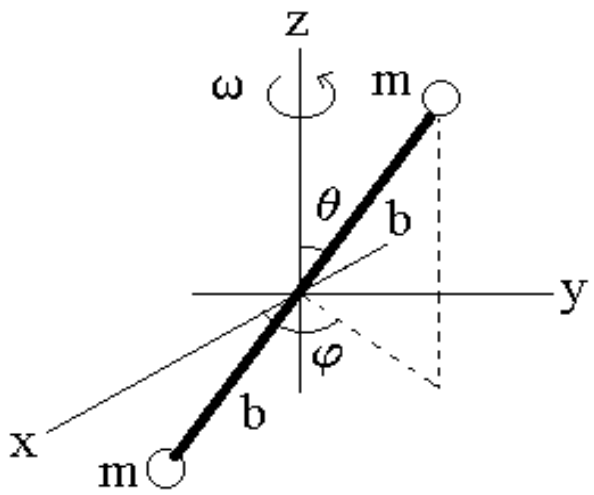


Figure 7.5: Dumbbell

where the unit vector $\hat{\mathsf{n}}$ defines the axis of rotation about which an angular rotation of angle α is performed according to the right-hand-rule.

A rigid body can be classified according to three categories. First, a rigid body is said to be a *spherical top* if its three principal moments of inertia are equal ($I_1 = I_2 = I_3$), i.e., the three roots of the cubic polynomial (7.16) are triply degenerate. Next, a rigid body is said to be a *symmetric top* if two of its principal moments of inertia are equal ($I_1 = I_2 \neq I_3$), i.e., I_3 is a single root and $I_1 = I_2$ are doubly-degenerate roots of the cubic polynomial (7.16). Lastly, when the three roots ($I_1 \neq I_2 \neq I_3$) are all single roots of the cubic polynomial (7.16), a rigid body is said to be an *asymmetric top*.

Before proceeding further, we consider the example of a dumbbell composed of two equal point masses m placed at the ends of a massless rod of total length $2\,b$ and rotating about the z-axis with angular frequency ω. Here, the positions of the two masses are expressed as

$$\mathbf{r}_\pm \;=\; \pm\, b\,[\sin\theta\,(\cos\varphi\,\hat{\mathsf{x}} \;+\; \sin\varphi\,\hat{\mathsf{y}}) \;+\; \cos\theta\,\hat{\mathsf{z}}\,]\,,$$

so that the CM inertia tensor is

$$\mathbf{I} \;=\; 2\,m\,b^2 \begin{pmatrix} 1-\cos^2\varphi\,\sin^2\theta & -\cos\varphi\,\sin\varphi\,\sin^2\theta & -\cos\varphi\,\cos\theta\,\sin\theta \\ -\cos\varphi\,\sin\varphi\,\sin^2\theta & 1-\sin^2\varphi\,\sin^2\theta & -\sin\varphi\,\cos\theta\,\sin\theta \\ -\cos\varphi\,\cos\theta\,\sin\theta & -\sin\varphi\,\cos\theta\,\sin\theta & 1-\cos^2\theta \end{pmatrix}. \quad (7.22)$$

After some tedious algebra, we find $\mathrm{Tr}(\mathbf{I}) = 4\,mb^2$, $\mathrm{Ad}(\mathbf{I}) = (2\,mb^2)^2$, and $\mathrm{Det}(\mathbf{I}) = 0$, and thus the cubic polynomial (7.16) has the single root $I_3 = 0$ and the double root $I_1 = I_2 = 2\,mb^2$, which makes the dumbbell a symmetric top.

The root $I_3 = 0$ clearly indicates that one of the three principal axes is the axis of symmetry of the dumbbell ($\hat{\mathsf{e}}_3 = \hat{r}$). The other two principal axes are located on the plane perpendicular to the symmetry axis (i.e., $\hat{\mathsf{e}}_1 = \hat{\theta}$ and $\hat{\mathsf{e}}_2 = \hat{\varphi}$). From these choices, we easily recover the rotation matrix

$$\mathsf{R} \;=\; \mathsf{R}_2(-\theta) \;\cdot\; \mathsf{R}_3(\varphi) \;=\; \begin{pmatrix} \cos\varphi\,\cos\theta & \sin\varphi\,\cos\theta & -\sin\theta \\ -\sin\varphi & \cos\varphi & 0 \\ \cos\varphi\,\sin\theta & \sin\varphi\,\sin\theta & \cos\theta \end{pmatrix},$$

so that, using the spherical coordinates (r, θ, φ), we find

$$\begin{aligned} \hat{\mathsf{e}}_1 &= \cos\theta\;(\cos\varphi\,\hat{\mathsf{x}} + \sin\varphi\,\hat{\mathsf{y}}) \;-\; \sin\theta\,\hat{\mathsf{z}} \;=\; \hat{\theta}, \\ \hat{\mathsf{e}}_2 &= -\;\sin\varphi\,\hat{\mathsf{x}} \;+\; \cos\varphi\,\hat{\mathsf{y}} \;=\; \hat{\varphi}, \\ \hat{\mathsf{e}}_3 &= \sin\theta\;(\cos\varphi\,\hat{\mathsf{x}} + \sin\varphi\,\hat{\mathsf{y}}) \;+\; \cos\theta\,\hat{\mathsf{z}} \;=\; \hat{r}. \end{aligned}$$

Indeed, the principal moment of inertia about the $\hat{r}$-axis is zero, while the principal moments of inertia about the perpendicular $\hat{\theta}$- and $\hat{\varphi}$-axes are equally given as $2\,mb^2$.

7.2 Angular Momentum

7.2.1 Euler Equations

The time derivative of the angular momentum $\mathbf{L} = \mathbf{I}\cdot\boldsymbol{\omega}$ in the fixed (LAB) frame is given as

$$\left(\frac{d\mathbf{L}}{dt}\right)_f \;=\; \left(\frac{d\mathbf{L}}{dt}\right)_r \;+\; \boldsymbol{\omega}\times\mathbf{L} \;=\; \mathbf{N},$$

where $\mathbf{N}$ represents the external torque applied to the system and $(d\mathbf{L}/dt)_r$ denotes the rate of change of $\mathbf{L}$ in the rotating frame. By choosing the body frame as the rotating frame, we find

$$\left(\frac{d\mathbf{L}}{dt}\right)_r \;=\; \mathbf{I}\cdot\dot{\boldsymbol{\omega}} \;=\; (I_1\,\dot{\omega}_1)\,\hat{\mathsf{e}}_1 \;+\; (I_2\,\dot{\omega}_2)\,\hat{\mathsf{e}}_2 \;+\; (I_3\,\dot{\omega}_3)\,\hat{\mathsf{e}}_3, \tag{7.23}$$

while

$$\boldsymbol{\omega}\times\mathbf{L} \;=\; -\,\hat{\mathsf{e}}_1\,\{\,\omega_2\,\omega_3\;(I_2 - I_3)\,\} \;-\; \hat{\mathsf{e}}_2\,\{\,\omega_3\,\omega_1\;(I_3 - I_1)\,\} \;-\; \hat{\mathsf{e}}_3\,\{\,\omega_1\,\omega_2\;(I_1 - I_2)\,\}. \tag{7.24}$$

Thus the time evolution of the angular momentum in the body frame of reference is described in terms of

$$\left.\begin{aligned} I_1\,\dot{\omega}_1 \;-\; \omega_2\,\omega_3\;(I_2 - I_3) &= N_1 \\ I_2\,\dot{\omega}_2 \;-\; \omega_3\,\omega_1\;(I_3 - I_1) &= N_2 \\ I_3\,\dot{\omega}_3 \;-\; \omega_1\,\omega_2\;(I_1 - I_2) &= N_3 \end{aligned}\right\}, \tag{7.25}$$

which are known as the Euler equations. Lastly, we note that the rate of change of the rotational kinetic energy (7.7) is expressed as

$$\frac{dK_{rot}}{dt} \;=\; \boldsymbol{\omega}\cdot\mathbf{I}\cdot\dot{\boldsymbol{\omega}} \;=\; \boldsymbol{\omega}\cdot\left(-\,\boldsymbol{\omega}\times\mathbf{L} \;+\; \mathbf{N}\right) \;=\; \mathbf{N}\cdot\boldsymbol{\omega}. \tag{7.26}$$

We note that in the absence of external torque ($\mathbf{N} = 0$), not only is kinetic energy conserved but also $L^2 = \sum_{i=1}^{3} (I_i\omega_i)^2$, as can be verified from Eq. (7.25). Note that, for a general function $F(\mathbf{L})$ of angular momentum and in the absence of external torque, we find

$$\frac{dF}{dt} \;=\; -\,\mathbf{L}\cdot\frac{\partial F}{\partial \mathbf{L}}\times\frac{\partial K_{rot}}{\partial \mathbf{L}} \;=\; -\,\frac{\partial F}{\partial \mathbf{L}}\cdot\boldsymbol{\omega}\times\mathbf{L},$$

and thus any function of $|\mathbf{L}|$ is a constant of the motion for rigid body dynamics.

7.2.2 Euler Equations for a Force-Free Symmetric Top

As an application of the Euler equations (7.25) we consider the case of the dynamics of a force-free symmetric top, for which $\mathbf{N} = 0$ and $I_1 = I_2 \neq I_3$. Accordingly, the Euler equations (7.25) become

$$\left.\begin{array}{lcl} I_1\,\dot{\omega}_1 &=& \omega_2\,\omega_3\;(I_1 - I_3) \\ I_1\,\dot{\omega}_2 &=& \omega_3\,\omega_1\;(I_3 - I_1) \\ I_3\,\dot{\omega}_3 &=& 0 \end{array}\right\}, \tag{7.27}$$

The last Euler equation states that if $I_3 \neq 0$, we have $\dot{\omega}_3 = 0$ or that ω_3 is a constant of motion. Next, after defining the precession frequency

$$\omega_p \;=\; \omega_3\left(\frac{I_3}{I_1} - 1\right), \tag{7.28}$$

which may be positive ($I_3 > I_1$) or negative ($I_3 < I_1$), the first two Euler equations yield

$$\dot{\omega}_1(t) \;=\; -\,\omega_p\,\omega_2(t) \quad\text{and}\quad \dot{\omega}_2(t) \;=\; \omega_p\,\omega_1(t). \tag{7.29}$$

The general solutions for $\omega_1(t)$ and $\omega_2(t)$ are

$$\omega_1(t) \;=\; \omega_0\;\cos(\omega_p t + \phi_0) \quad\text{and}\quad \omega_2(t) \;=\; \omega_0\;\sin(\omega_p t + \phi_0), \tag{7.30}$$

where ω_0 is a constant and ϕ_0 is an initial phase associated with initial conditions for $\omega_1(t)$ and $\omega_2(t)$. Since ω_3 and $\omega_0^2 = \omega_1^2(t) + \omega_2^2(t)$ are constant, then the magnitude of the angular velocity $\boldsymbol{\omega}$,

$$\omega \;=\; \sqrt{\omega_1^2 + \omega_2^2 + \omega_3^2},$$

is also a constant. Thus the angle α between $\boldsymbol{\omega}$ and $\widehat{\mathsf{e}}_3$ is constant, with

$$\omega_3 \;=\; \omega\,\cos\alpha \quad\text{and}\quad \sqrt{\omega_1^2 + \omega_2^2} \;=\; \omega_0 \;=\; \omega\,\sin\alpha.$$

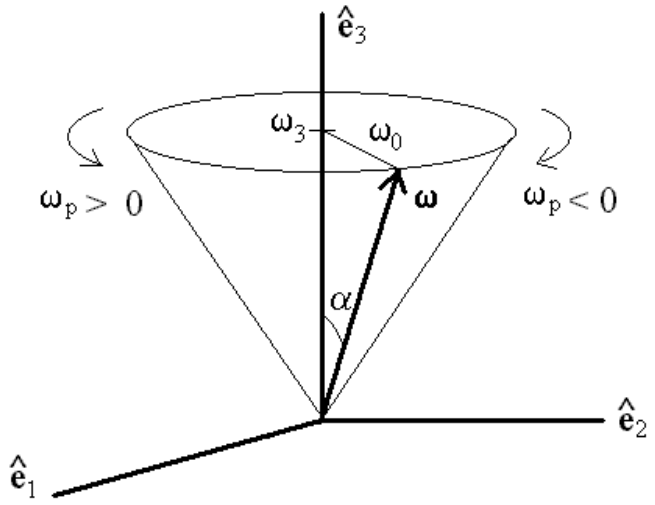


Figure 7.6: Body cone

Since the magnitude of $\boldsymbol{\omega}$ is also constant, the $\boldsymbol{\omega}$-dynamics simply involves a constant rotation with frequency ω_3 and a precession motion of $\boldsymbol{\omega}$ about the $\hat{\mathsf{e}}_3$-axis with a precession frequency ω_p; as a result of precession, the vector $\boldsymbol{\omega}$ spans the *body* cone with $\omega_p > 0$ if $I_3 > I_1$ (for a pancake-shaped or oblate symmetric top) or $\omega_p < 0$ if $I_3 < I_1$ (for a cigar-shaped or prolate symmetric top).

For example, to a good approximation, Earth is an oblate spheroid with

$$I_1 \;=\; \frac{1}{5}\,M\,\left(a^2+c^2\right) \;=\; I_2 \;\text{ and }\; I_3 \;=\; \frac{2}{5}\,M\,a^2 \;>\; I_1,$$

where $2\,c = 12,714$ km is the Pole-to-Pole distance and $2\,a = 12,756$ km is the equatorial diameter, so that

$$\frac{I_3}{I_1} \;-\; 1 \;=\; \frac{a^2-c^2}{a^2+c^2} \;=\; 0.003298... \;=\; \epsilon.$$

The precession frequency (7.28) of the rotation axis of Earth is, therefore, $\omega_p = \epsilon\,\omega_3$, where $\omega_3 = 2\pi$ rad/day is the rotation frequency of the Earth, so that the precession motion repeats itself every ϵ^{-1} days or 303 days; the actual period is 430 days and the difference is partially due to the non-rigidity of Earth and the fact that the Earth is not a pure oblate spheroid. A slower precession motion of approximately 26,000 years is introduced by the combined gravitational effect of the Sun and the Moon on one hand, and the fact that the Earth's rotation axis is at an angle 23.5^o to the Ecliptic plane (on which most planets move).

The fact that the symmetric top is force-free implies that its rotational kinetic energy is constant [see Eq. (7.26)] and, hence, $\mathbf{L}\cdot\boldsymbol{\omega}$ is constant while $\boldsymbol{\omega}\times\mathbf{L}\cdot\hat{\mathsf{e}}_3 = 0$ according to

Eq. (7.24). Since $\mathbf{L}$ itself is constant in magnitude and direction in the LAB (or fixed) frame, we may choose the $\hat{\mathbf{z}}$-axis to be along $\mathbf{L}$ (i.e., $\mathbf{L} = \ell\,\hat{\mathbf{z}}$). If at a given instant, $\omega_1 = 0$, then $\omega_2 = \omega_0 = \omega\,\sin\alpha$ and $\omega_3 = \omega\,\cos\alpha$. Likewise, we may write $L_1 = I_1\,\omega_1 = 0$, and

$$\begin{aligned} L_2 &= I_2\,\omega_2 = I_1\,\omega\,\sin\alpha = \ell\,\sin\theta, \\ L_3 &= I_3\,\omega_3 = I_3\,\omega\,\cos\alpha = \ell\,\cos\theta, \end{aligned}$$

where $\mathbf{L}\cdot\boldsymbol{\omega} = \ell\,\omega\,\cos\theta$, with θ represents the *space*-cone angle. From these equations, we find the relation between the body-cone angle α and the space-cone angle θ to be

$$\tan\theta = \left(\frac{I_1}{I_3}\right)\,\tan\alpha, \tag{7.31}$$

which shows that $\theta > \alpha$ for $I_3 < I_1$ and $\theta < \alpha$ for $I_3 > I_1$.

7.2.3 Euler Equations for a Force-Free Asymmetric Top

We now consider the general case of an asymmetric top moving under force-free conditions. To facilitate our discussion, we assume that $I_1 > I_2 > I_3$ and thus Euler's equations (7.25) for a force-free asymmetric top are

$$\left.\begin{aligned} I_1\,\dot{\omega}_1 &= \omega_2\,\omega_3\,\,(I_2 - I_3) \\ I_2\,\dot{\omega}_2 &= -\,\omega_3\,\omega_1\,\,(I_1 - I_3) \\ I_3\,\dot{\omega}_3 &= \omega_1\,\omega_2\,\,(I_1 - I_2) \end{aligned}\right\}. \tag{7.32}$$

As previously mentioned, the Euler equations (7.32) have two constants of the motion: kinetic energy

$$K = \frac{1}{2}\left(I_1\,\omega_1^2 + I_2\,\omega_2^2 + I_3\,\omega_3^2\right), \tag{7.33}$$

and the squared magnitude of the angular momentum

$$L^2 = I_1^2\,\omega_1^2 + I_2^2\,\omega_2^2 + I_3^3\,\omega_3^2. \tag{7.34}$$

Figure 7.7 shows the numerical solution of the Euler equations (7.32) subject to the initial condition $(\omega_{10}, \omega_{20}, \omega_{30}) = (2, 0, 1)$ for different values of the ratios I_1/I_3 and I_2/I_3. Note that in the limit $I_1 = I_2$ (corresponding to a symmetric top), the top evolves solely on the (ω_1, ω_2)-plane at constant ω_3. As I_1 increases from I_2, the asymmetric top exhibits doubly-periodic behavior in the full $(\omega_1, \omega_2, \omega_3)$-space until the motion becomes restricted to the (ω_2, ω_3)-plane in the limit $I_1 \gg I_2$. One also clearly notes the existence of a separatrix which appears as I_1 reaches the critical value

$$I_{1c} = \frac{I_2}{2} + \sqrt{\frac{I_2^2}{4} + I_3\,(I_2 - I_3)\left(\frac{\omega_{30}}{\omega_{10}}\right)^2},$$

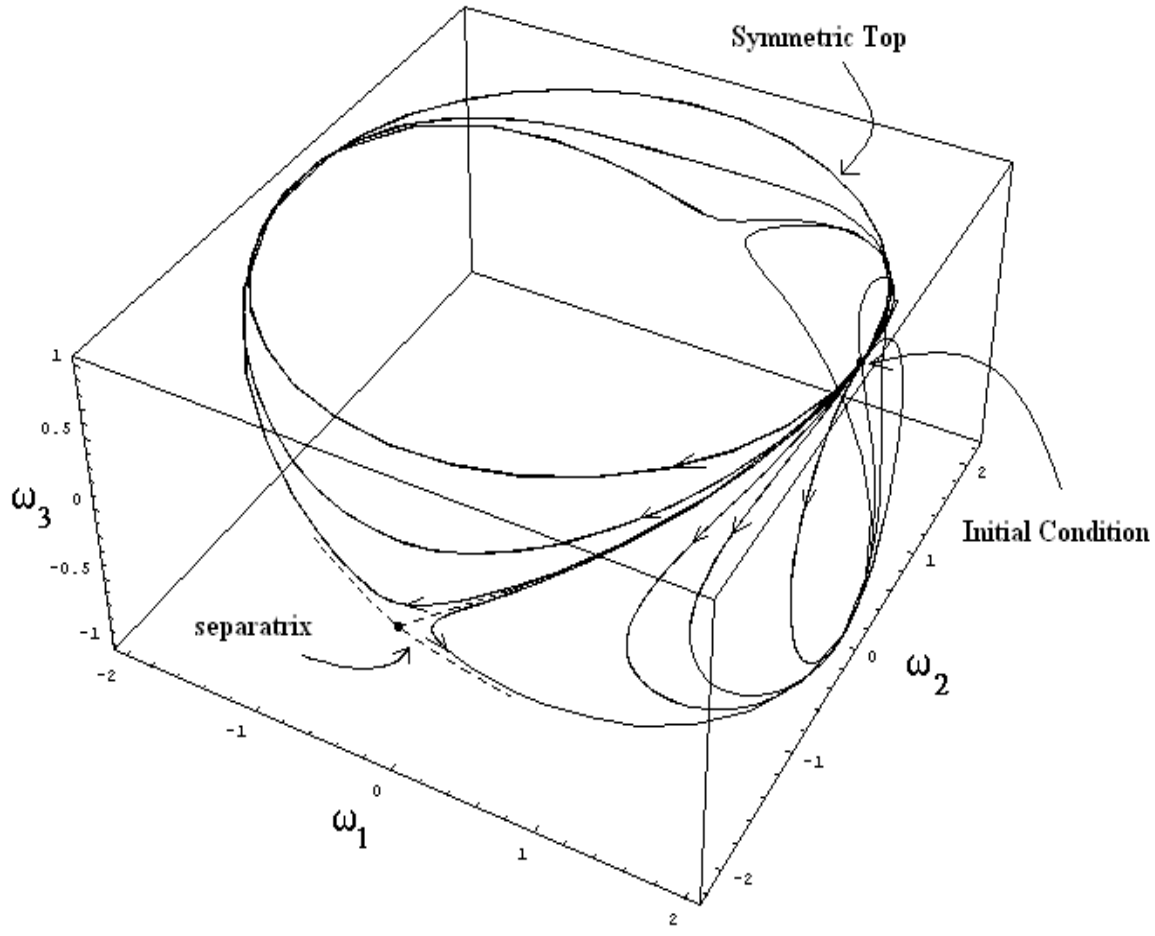


Figure 7.7: Orbits of an asymmetric top

at constant I_2 and I_3 and given initial conditions $(\omega_{10}, \omega_{20}, \omega_{30})$.

We note that the existence of two constants of the motion, Eqs. (7.33) and (7.34), for the three Euler equations (7.32) means that we may express the Euler equations in terms of a single equation. For this purpose, we introduce the constants

$$\sigma \;=\; 2\,I_1K \;-\; L^2 \quad \text{and} \quad \rho \;=\; L^2 \;-\; 2\,I_3K,$$

from which we obtain expressions for ω_1 (taken here to be negative) and ω_3 in terms of ω_2:

$$\omega_1 \;=\; -\sqrt{\frac{\rho \;-\; I_2(I_2 - I_3)\,\omega_2^2}{I_1\,(I_1 - I_3)}} \quad \text{and} \quad \omega_3 \;=\; \sqrt{\frac{\sigma \;-\; I_2(I_1 - I_2)\,\omega_2^2}{I_3\,(I_1 - I_3)}}. \tag{7.35}$$

When we substitute these expressions in the Euler equation for ω_2, we easily obtain

$$\dot{\omega}_2 \;=\; \alpha\,\sqrt{(\Omega_1^2 \;-\; \omega_2^2)\,(\Omega_3^2 \;-\; \omega_2^2)}, \tag{7.36}$$

where α is a positive dimensionless constant defined as

$$\alpha \;=\; \sqrt{\left(1 - \frac{I_2}{I_1}\right)\,\left(\frac{I_2}{I_3} - 1\right)},$$

while the constant frequencies Ω_1 and Ω_3 are defined as

$$\Omega_1^2 \;=\; \frac{2\,I_1K \;-\; L^2}{I_2(I_1 - I_2)} \quad \text{and} \quad \Omega_3^2 \;=\; \frac{L^2 \;-\; 2\,I_3K}{I_2(I_2 - I_3)}.$$

We immediately note that the evolution of ω_2 is characterized by the two frequencies Ω_1 and Ω_3, which also represent the turning points at which $\dot{\omega}_2$ vanishes. Next, by introducing a dimensionless frequency $u \;=\; \omega_2/\Omega_1$ (here, we assume that $\Omega_1 \geq \Omega_3$) and a dimensionless time $\tau = \alpha\Omega_3\,t$, the Euler equation (7.36) can now be integrated to yield

$$\tau \;=\; \int_0^u \frac{ds}{\sqrt{(1 - s^2)\,(1 - k^2\,s^2)}} \;=\; \int_0^{\Theta(u)} \frac{d\theta}{\sqrt{1 - k^2\,\sin^2\theta}}, \tag{7.37}$$

where $k^2 = \Omega_1^2/\Omega_3^2 \geq 1$, $\Theta(u) = \arcsin u$, and we assume that $\omega_2(t = 0) = 0$; compare Eq. (7.37) with Eq. (3.25).

Lastly, the turning points for ω_2 are now represented in terms of turning points for u as $\omega_2 = \Omega_1 \;\;\rightarrow\;\; u = 1$ and $\omega_2 = \Omega_3 \;\;\rightarrow\;\; u = k^{-1} = \Omega_3/\Omega_1 \leq 1$. Lastly, note that the separatrix solution of the force-free asymmetric top (see Figure 7.7) corresponding to $I_1 = I_{1c}$ is associated with $k = 1$ (i.e., $\Omega_1 = \Omega_3$).

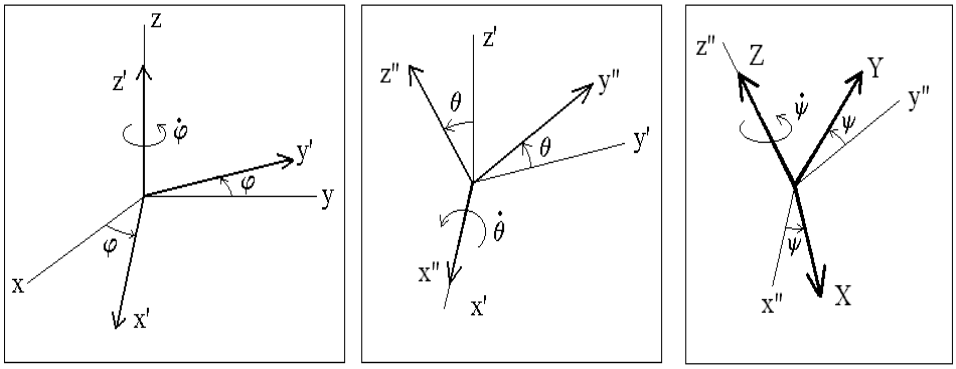


Figure 7.8: Euler angles

7.3 Symmetric Top with One Fixed Point

7.3.1 Eulerian Angles as generalized Lagrangian Coordinates

To describe the physical state of a rotating object with principal moments of inertia (I_1, I_2, I_3), we need the three Eulerian angles (φ, θ, ψ) in the body frame of reference (see Figure 7.8). The Eulerian angle φ is associated with the rotation of the fixed-frame unit vectors $(\hat{\mathsf{x}}, \hat{\mathsf{y}}, \hat{\mathsf{z}})$ about the z-axis. We thus obtain the new unit vectors $(\hat{\mathsf{x}}', \hat{\mathsf{y}}', \hat{\mathsf{z}}')$ defined as

$$\begin{pmatrix} \hat{\mathsf{x}}' \\ \hat{\mathsf{y}}' \\ \hat{\mathsf{z}}' \end{pmatrix} = \overbrace{\begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}}^{=\mathsf{R}_3(\varphi)} \cdot \begin{pmatrix} \hat{\mathsf{x}} \\ \hat{\mathsf{y}} \\ \hat{\mathsf{z}} \end{pmatrix} \tag{7.38}$$

The rotation matrix $\mathsf{R}_3(\varphi)$ has the following properties associated with a general rotation matrix $\mathsf{R}_i(\alpha)$, where a rotation of axes about the x^i-axis is performed through an arbitrary angle α. First, the matrix $\mathsf{R}_i(-\alpha)$ is the inverse matrix of $\mathsf{R}_i(\alpha)$, i.e.,

$$\mathsf{R}_i(-\alpha) \cdot \mathsf{R}_i(\alpha) \;=\; \underline{\mathbf{1}} \;=\; \mathsf{R}_i(\alpha) \cdot \mathsf{R}_i(-\alpha).$$

Next, the determinant of $\mathsf{R}_i(\alpha)$ is

$$\mathrm{Det}[\mathsf{R}_i(\alpha)] \;=\; +1.$$

Lastly, the eigenvalues of $\mathsf{R}_i(\alpha)$ are $+1$, $e^{i\alpha}$, and $e^{-i\alpha}$.

The Eulerian angle θ is associated with the rotation of the unit vectors $(\hat{\mathsf{x}}', \hat{\mathsf{y}}', \hat{\mathsf{z}}')$ about the x'-axis. We thus obtain the new unit vectors $(\hat{\mathsf{x}}'', \hat{\mathsf{y}}'', \hat{\mathsf{z}}'')$ defined as

$$\begin{pmatrix} \hat{\mathsf{x}}'' \\ \hat{\mathsf{y}}'' \\ \hat{\mathsf{z}}'' \end{pmatrix} = \overbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}}^{=\mathsf{R}_1(\theta)} \cdot \begin{pmatrix} \hat{\mathsf{x}}' \\ \hat{\mathsf{y}}' \\ \hat{\mathsf{z}}' \end{pmatrix} \tag{7.39}$$

The Eulerian angle ψ is associated with the rotation of the unit vectors $(\widehat{\mathsf{x}}'', \widehat{\mathsf{y}}'', \widehat{\mathsf{z}}'')$ about the z''-axis. We thus obtain the body-frame unit vectors $(\widehat{\mathsf{e}}_1, \widehat{\mathsf{e}}_2, \widehat{\mathsf{e}}_3)$ defined as

$$\begin{pmatrix} \widehat{\mathsf{e}}_1 \\ \widehat{\mathsf{e}}_2 \\ \widehat{\mathsf{e}}_3 \end{pmatrix} = \overbrace{\begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}}^{=\mathsf{R}_3(\psi)} \cdot \begin{pmatrix} \widehat{\mathsf{x}}'' \\ \widehat{\mathsf{y}}'' \\ \widehat{\mathsf{z}}'' \end{pmatrix} \tag{7.40}$$

Hence, the relation between the fixed-frame unit vectors $(\widehat{\mathsf{x}}, \widehat{\mathsf{y}}, \widehat{\mathsf{z}})$ and the body-frame unit vectors $(\widehat{\mathsf{e}}_1, \widehat{\mathsf{e}}_2, \widehat{\mathsf{e}}_3)$ involves the matrix $\mathsf{R} = \mathsf{R}_3(\psi) \cdot \mathsf{R}_1(\theta) \cdot \mathsf{R}_3(\varphi)$, such that $\widehat{\mathsf{e}}_i = R_{ij}\, \widehat{\mathsf{x}}^j$, or

$$\left.\begin{aligned} \widehat{\mathsf{e}}_1 &= \cos\psi\, \widehat{\perp} + \sin\psi\,(\cos\theta\, \widehat{\varphi} + \sin\theta\, \widehat{\mathsf{z}}) \\ \widehat{\mathsf{e}}_2 &= -\sin\psi\, \widehat{\perp} + \cos\psi\,(\cos\theta\, \widehat{\varphi} + \sin\theta\, \widehat{\mathsf{z}}) \\ \widehat{\mathsf{e}}_3 &= -\sin\theta\, \widehat{\varphi} + \cos\theta\, \widehat{\mathsf{z}} \end{aligned}\right\}, \tag{7.41}$$

where $\widehat{\varphi} = -\sin\varphi\, \widehat{\mathsf{x}} + \cos\varphi\, \widehat{\mathsf{y}}$ and $\widehat{\perp} = \cos\varphi\, \widehat{\mathsf{x}} + \sin\varphi\, \widehat{\mathsf{y}} = \widehat{\varphi} \times \widehat{\mathsf{z}}$.

7.3.2 Angular Velocity in terms of Eulerian Angles

The angular velocity $\boldsymbol{\omega}$ represented in the three Figures above is expressed as

$$\boldsymbol{\omega} = \dot{\varphi}\, \widehat{\mathsf{z}} + \dot{\theta}\, \widehat{\mathsf{x}}' + \dot{\psi}\, \widehat{\mathsf{e}}_3.$$

The unit vectors $\widehat{\mathsf{z}}$ and $\widehat{\mathsf{x}}'$ are written in terms of the body-frame unit vectors $(\widehat{\mathsf{e}}_1, \widehat{\mathsf{e}}_2, \widehat{\mathsf{e}}_3)$ as

$$\begin{aligned} \widehat{\mathsf{z}} &= \sin\theta\, (\sin\psi\, \widehat{\mathsf{e}}_1 + \cos\psi\, \widehat{\mathsf{e}}_2) + \cos\theta\, \widehat{\mathsf{e}}_3, \\ \widehat{\mathsf{x}}' &= \widehat{\mathsf{x}}'' = \cos\psi\, \widehat{\mathsf{e}}_1 - \sin\psi\, \widehat{\mathsf{e}}_2. \end{aligned}$$

The angular velocity can, therefore, be written exclusively in the body frame of reference in terms of the Euler basis vectors (7.41) as

$$\boldsymbol{\omega} = \omega_1\, \widehat{\mathsf{e}}_1 + \omega_2\, \widehat{\mathsf{e}}_2 + \omega_3\, \widehat{\mathsf{e}}_3, \tag{7.42}$$

where the body-frame angular frequencies are

$$\left.\begin{aligned} \omega_1 &= \dot{\varphi}\, \sin\theta\, \sin\psi + \dot{\theta}\, \cos\psi \\ \omega_2 &= \dot{\varphi}\, \sin\theta\, \cos\psi - \dot{\theta}\, \sin\psi \\ \omega_3 &= \dot{\psi} + \dot{\varphi}\, \cos\theta \end{aligned}\right\}. \tag{7.43}$$

Note that all three frequencies are independent of φ (i.e., $\partial\omega_i/\partial\varphi = 0$), while derivatives with respect to ψ and $\dot{\psi}$ are

$$\frac{\partial\omega_1}{\partial\psi} = \omega_2, \quad \frac{\partial\omega_2}{\partial\psi} = -\omega_1, \quad \text{and} \quad \frac{\partial\omega_3}{\partial\psi} = 0,$$

and

$$\frac{\partial \omega_1}{\partial \dot{\psi}} \;=\; 0 \;=\; \frac{\partial \omega_2}{\partial \dot{\psi}} \quad \text{and} \quad \frac{\partial \omega_3}{\partial \dot{\psi}} \;=\; 1.$$

The relations (7.43) can be inverted to yield

$$\begin{aligned} \dot{\varphi} &= \csc\theta\,(\sin\psi\,\omega_1 \;+\; \cos\psi\,\omega_2), \\ \dot{\theta} &= \cos\psi\,\omega_1 \;-\; \sin\psi\,\omega_2, \\ \dot{\psi} &= \omega_3 \;-\; \cot\theta\,(\sin\psi\,\omega_1 \;+\; \cos\psi\,\omega_2). \end{aligned}$$

7.3.3 Rotational Kinetic Energy of a Symmetric Top

The rotational kinetic energy (7.7) for a symmetric top can be written as

$$K_{rot} \;=\; \frac{1}{2}\,\left\{\, I_3\,\omega_3^2 \;+\; I_1\left(\omega_1^2 \;+\; \omega_2^2\right)\,\right\},$$

or explicitly in terms of the Eulerian angles (φ, θ, ψ) and their time derivatives $(\dot{\varphi}, \dot{\theta}, \dot{\psi})$ as

$$K_{rot} \;=\; \frac{1}{2}\,\left\{\, I_3\left(\dot{\psi} \;+\; \dot{\varphi}\,\cos\theta\right)^2 \;+\; I_1\left(\dot{\theta}^2 \;+\; \dot{\varphi}^2\,\sin^2\theta\right)\,\right\}. \tag{7.44}$$

We now briefly return to the case of the force-free symmetric top for which the Lagrangian is simply $L(\theta, \dot{\theta};\; \dot{\varphi}, \dot{\psi}) = K_{rot}$. Since φ and ψ are ignorable coordinates, i.e., the force-free Lagrangian (7.44) is independent of φ and ψ, their canonical angular momenta

$$p_\varphi \;=\; \frac{\partial L}{\partial \dot{\varphi}} \;=\; I_3\,(\dot{\psi} + \dot{\varphi}\,\cos\theta)\,\cos\theta \;+\; I_1\,\sin^2\theta\,\dot{\varphi}, \tag{7.45}$$

$$p_\psi \;=\; \frac{\partial L}{\partial \dot{\psi}} \;=\; I_3\,(\dot{\psi} + \dot{\varphi}\,\cos\theta) \;=\; I_3\,\omega_3 \tag{7.46}$$

are constants of the motion. By inverting these relations, we obtain

$$\dot{\varphi} \;=\; \frac{p_\varphi - p_\psi\,\cos\theta}{I_1\,\sin^2\theta} \quad \text{and} \quad \dot{\psi} \;=\; \omega_3 \;-\; \frac{(p_\varphi - p_\psi\,\cos\theta)\,\cos\theta}{I_1\,\sin^2\theta}, \tag{7.47}$$

and the rotational kinetic energy (7.44) becomes

$$K_{rot} \;=\; \frac{1}{2}\left\{\, I_1\,\dot{\theta}^2 \;+\; I_3\,\omega_3^2 \;+\; \frac{(p_\varphi - p_\psi\,\cos\theta)^2}{I_1\,\sin^2\theta}\,\right\}. \tag{7.48}$$

The motion of a force-free symmetric top can now be described in terms of solutions of the Euler-Lagrange equation for the Eulerian angle θ:

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) \;=\; I_1\,\ddot{\theta} \;&=\; \frac{\partial L}{\partial \theta} \;=\; \dot{\varphi}\,\sin\theta\,(I_1\,\cos\theta\,\dot{\varphi} \;-\; p_\psi) \\ &=\; -\,\frac{(p_\varphi - p_\psi\,\cos\theta)}{I_1\,\sin\theta}\,\frac{(p_\psi - p_\varphi\,\cos\theta)}{\sin^2\theta}. \end{aligned} \tag{7.49}$$

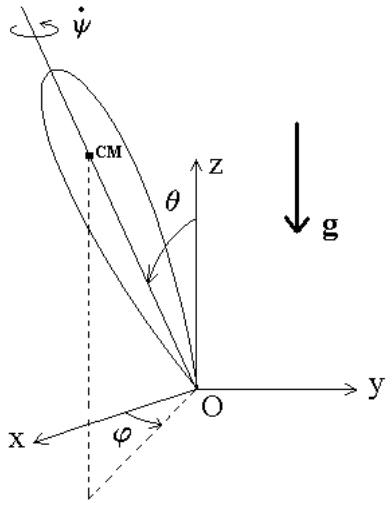


Figure 7.9: Symmetric top with one fixed point

Once $\theta(t)$ is solved for given values of the principal moments of inertia $I_1 = I_2$ and I_3 and the invariant canonical angular momenta p_φ and p_ψ, the functions $\varphi(t)$ and $\psi(t)$ are determined from the time integration of Eqs. (7.47).

7.3.4 Lagrangian Dynamics of a Symmetric Top with One Fixed Point

We now consider the case of a spinning symmetric top of mass M and principal moments of inertia ($I_1 = I_2 \neq I_3$) with one fixed point O moving in a gravitational field with constant acceleration g. The rotational kinetic energy of the symmetric top is given by Eq. (7.44) while the potential energy for the case of a symmetric top with one fixed point is

$$U(\theta) \;=\; Mgh\,\cos\theta, \tag{7.50}$$

where h is the distance from the fixed point O and the center of mass (CM) of the symmetric top. The Lagrangian for the symmetric top with one fixed point is

$$L(\theta,\dot\theta;\,\dot\varphi,\dot\psi) \;=\; \frac{1}{2}\,\left\{\, I_3\left(\dot\psi \;+\; \dot\varphi\,\cos\theta\right)^2 \;+\; I_1\left(\dot\theta^2 \;+\; \dot\varphi^2\,\sin^2\theta\right)\,\right\} \;-\; Mgh\,\cos\theta. \tag{7.51}$$

A normalized form of the Euler equations for the symmetric top with one fixed point (also known as the *heavy* symmetric top) is expressed as

$$\varphi' \;=\; \frac{(b-\cos\theta)}{\sin^2\theta} \quad\text{and}\quad \theta'' \;=\; a\,\sin\theta \;+\; \frac{(1-b\,\cos\theta)(b-\cos\theta)}{\sin^3\theta}, \tag{7.52}$$

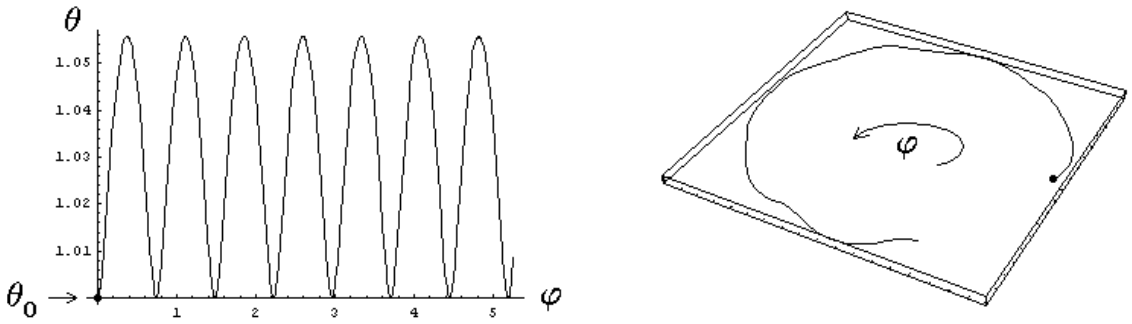


Figure 7.10: Orbits of heavy top – Case I

where time has been rescaled such that $(\cdots)' = (I_1/p_\psi)\,(\cdots)^{\bullet}$ and the two parameters a and b are defined as

$$a \;=\; \frac{Mgh\,I_1}{p_\psi^2} \quad \text{and} \quad b \;=\; \frac{p_\varphi}{p_\psi}.$$

The normalized heavy-top equations (7.52) have been integrated for the initial conditions $(\theta_0, \theta_0'; \varphi_0) = (1.0, 0.0; 0.0)$. The three Figures shown below correspond to three different cases (I, II, and III) for fixed value of a (here, $a = 0.1$), which exhibit the possibility of azimuthal reversal when φ' changes sign for different values of $b = p_\varphi/p_\psi$; the azimuthal precession motion is called *nutation*.

The Figures on the left show the normalized heavy-top solutions in the (φ, θ)-plane while the Figures on the right show the spherical projection of the normalized heavy-top solutions $(\theta, \varphi) \to (\sin\theta\,\cos\varphi,\ \sin\theta\,\sin\varphi,\ \cos\theta)$, where the initial condition is denoted by a dot ($\bullet$). In Case I ($b > \cos\theta_0$), the azimuthal velocity φ' never changes sign and azimuthal precession occurs monotonically. In Case II ($b = \cos\theta_0$), the azimuthal velocity φ' vanishes at $\theta = \theta_0$ (where θ' also vanishes) and the heavy symmetric top exhibits a *cusp* at $\theta = \theta_0$. In Case III ($b < \cos\theta_0$), the azimuthal velocity φ' vanishes for $\theta > \theta_0$ and the heavy symmetric top exhibits a phase of *retrograde* motion. Since the Lagrangian (7.51) is independent of the Eulerian angles φ and ψ, the canonical angular momenta p_φ and p_ψ, respectively, are constants of the motion. The solution for $\theta(t)$ is then most easily obtained by considering the energy equation

$$E \;=\; \frac{1}{2}\left\{\, I_1\,\dot\theta^2 \;+\; I_3\,\omega_3^2 \;+\; \frac{(p_\varphi - p_\psi\,\cos\theta)^2}{I_1\,\sin^2\theta} \,\right\} \;+\; Mgh\,\cos\theta, \tag{7.53}$$

where p_φ and $p_\psi = I_3\,\omega_3$ are constants of the motion. Since the total energy E is itself a constant of the motion, we may define a new energy constant

$$E' \;=\; E \;-\; \frac{1}{2}\,I_3\,\omega_3^2,$$

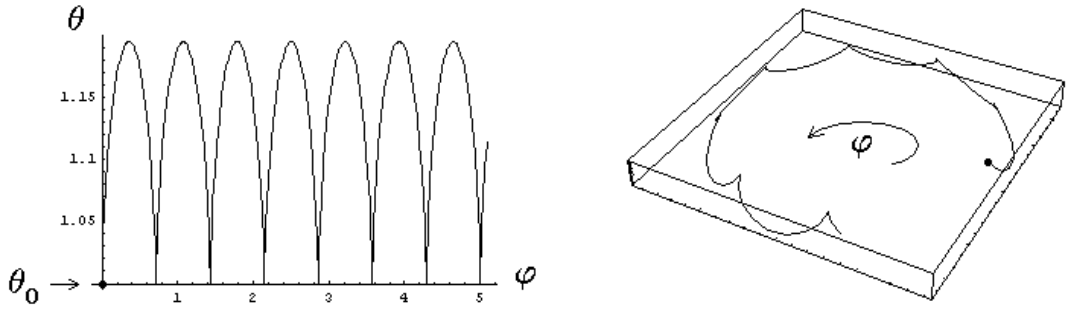


Figure 7.11: Orbits of heavy top – Case II

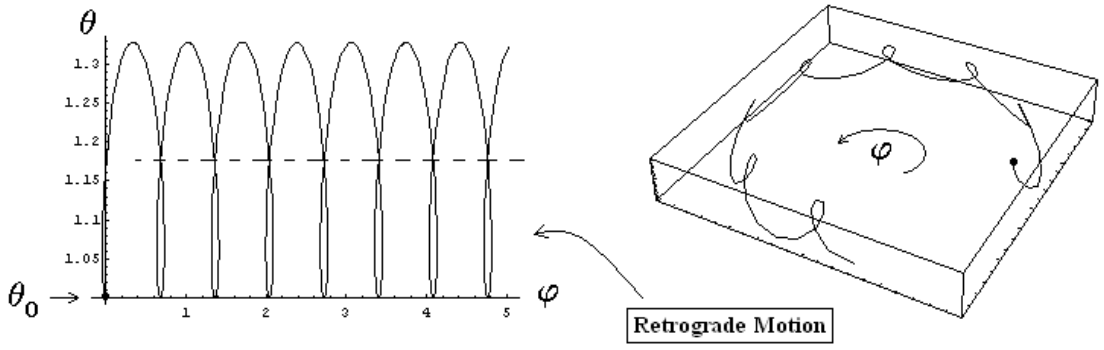


Figure 7.12: Orbits of heavy top – Case III

and an effective potential energy

$$V(\theta) \;=\; \frac{(p_\varphi - p_\psi\,\cos\theta)^2}{2I_1\,\sin^2\theta} \;+\; Mgh\,\cos\theta, \tag{7.54}$$

so that Eq. (7.53) becomes

$$E' \;=\; \frac{1}{2}\,I_1\,\dot\theta^2(t) \;+\; V(\theta), \tag{7.55}$$

which can be *formally solved* as

$$t(\theta) \;=\; \pm\int \frac{d\theta}{\sqrt{(2/I_1)\,[E' - V(\theta)]}}. \tag{7.56}$$

Note that turning points θ_{tp} are again defined as roots of the equation $E' = V(\theta)$.

A simpler formulation for this problem is obtained as follows. First, we define the following quantities

$$\Omega^2 \;=\; \frac{2\,Mgh}{I_1},\;\; \epsilon \;=\; \frac{2\,E'}{I_1\Omega^2} \;=\; \frac{E'}{Mgh},\;\; \alpha \;=\; \frac{p_\varphi}{I_1\Omega},\;\; \text{and}\;\; \beta \;=\; \frac{p_\psi}{I_1\Omega}, \tag{7.57}$$

so that Eq. (7.56) becomes

$$\tau(u) \;=\; \pm\int \frac{du}{\sqrt{(1-u^2)(\epsilon-u) - (\alpha-\beta\,u)^2}} \;=\; \pm\int \frac{du}{\sqrt{(1-u^2)[\epsilon - W(u)]}}, \tag{7.58}$$

where $\tau(u) = \Omega\,t(u)$, $u = \cos\theta$, and the energy equation (7.55) becomes

$$\epsilon \;=\; \frac{1}{(1-u^2)}\left[\left(\frac{du}{d\tau}\right)^2 \;+\; (\alpha \;-\; \beta\,u)^2\right] \;+\; u \;=\; (1-u^2)^{-1}\left(\frac{du}{d\tau}\right)^2 \;+\; W(u), \tag{7.59}$$

where the effective potential is

$$W(u) \;=\; u \;+\; \frac{(\alpha-\beta\,u)^2}{(1-u^2)}.$$

We note that the effective potential $W(u)$ is infinite at $u = \pm 1$ and has a single minimum at $u = u_0$ (or $\theta = \theta_0$) defined by the quartic equation

$$W'(u_0) \;=\; 1 \;+\; 2u_0\left(\frac{\alpha-\beta\,u_0}{1-u_0^2}\right)^2 \;-\; 2\beta\left(\frac{\alpha-\beta\,u_0}{1-u_0^2}\right) \;=\; 0. \tag{7.60}$$

This equation has four roots: two roots are complex roots, a third root is always greater than one for $\alpha > 0$ and $\beta > 0$ (which is unphysical since $u = \cos\theta \le 1$), while the fourth root is less than one for $\alpha > 0$ and $\beta > 0$; hence, this root is the only physical root corresponding to a single minimum for the effective potential $W(u)$ (see Figure 7.13). Note

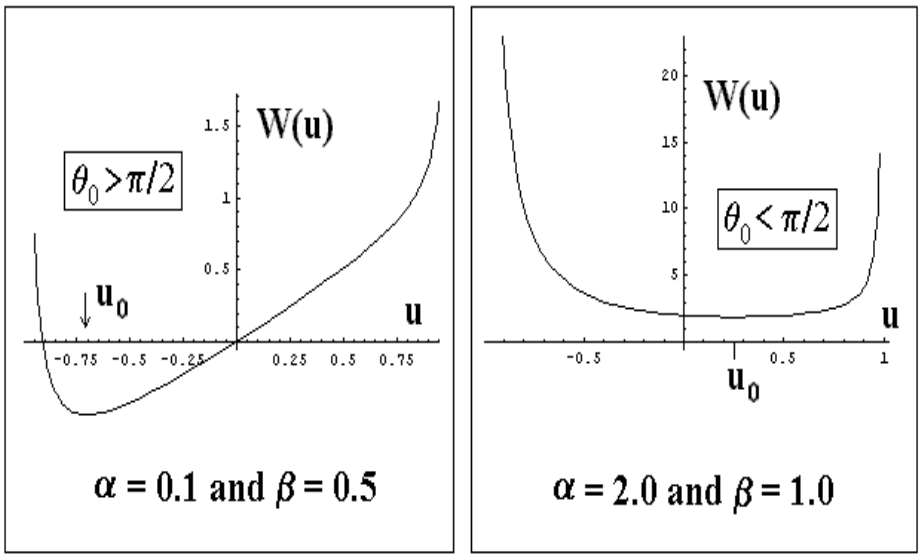


Figure 7.13: Effective potential for the heavy top

how the linear gravitational-potential term u is apparent at low values of α and β.

We first investigate the motion of the symmetric top at the minimum angle θ_0 for which $\epsilon = W(u_0)$ and $\dot{u}(u_0) = 0$. For this purpose, we note that when the dimensionless azimuthal frequency

$$\frac{d\varphi}{d\tau} \;=\; \frac{\alpha - \beta\, u}{1 - u^2} \;=\; \nu(u)$$

is inserted in Eq. (7.60), we obtain the quadratic equation $1 + 2u_0\,\nu_0^2 - 2\beta\,\nu_0 = 0$, which has two solutions for $\nu_0 = \nu(u_0)$:

$$\nu(u_0) \;=\; \frac{\beta}{2u_0}\left(1 \;\pm\; \sqrt{1 \;-\; \frac{2\,u_0}{\beta^2}}\right).$$

Here, we further note that these solutions require that the radicand be positive, i.e.,

$$\beta^2 \;>\; 2\,u_0 \quad \text{or} \quad I_3\,\omega_3 \geq I_1\Omega\,\sqrt{2\,u_0},$$

if $u_0 \geq 0$ (or $\theta_0 \leq \pi/2$); no conditions are applied to ω_3 for the case $u_0 < 0$ (or $\theta_0 > \pi/2$) since the radicand is strictly positive in this case.

Hence, the precession frequency $\dot{\varphi}_0 = \nu(u_0)\,\Omega$ at $\theta = \theta_0$ has a slow component and a fast component

$$(\dot{\varphi}_0)_{slow} \;=\; \frac{I_3\omega_3}{2\,I_1\,\cos\theta_0}\left[1 \;-\; \sqrt{1 \;-\; 2\left(\frac{I_1\Omega}{I_3\omega_3}\right)^2\cos\theta_0}\right],$$

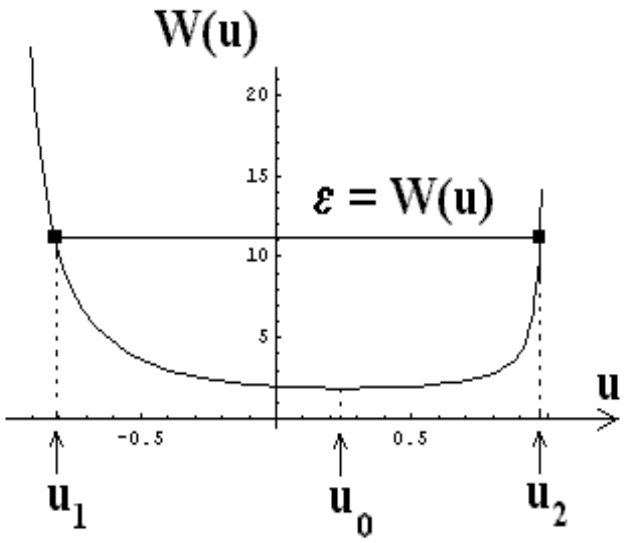


Figure 7.14: Turning-point roots

$$(\dot{\varphi}_0)_{fast} \;=\; \frac{I_3\omega_3}{2\,I_1\,\cos\theta_0}\,\left[\,1\;+\;\sqrt{1\;-\;2\left(\frac{I_1\Omega}{I_3\omega_3}\right)^2\cos\theta_0}\,\right].$$

We note that for $\theta_0 < \pi/2$ (or $\cos\theta_0 > 0$) the two precession frequencies $(\dot{\varphi}_0)_{slow}$ and $(\dot{\varphi}_0)_{fast}$ have the same sign while for $\theta_0 > \pi/2$ (or $\cos\theta_0 < 0$) the two precession frequencies have opposite signs $(\dot{\varphi}_0)_{slow} < 0$ and $(\dot{\varphi}_0)_{fast} > 0$.

Next, we investigate the case with two turning points $u_1 < u_0$ and $u_2 > u_1$ (or $\theta_1 > \theta_2$) where $\epsilon = W(u)$ (see Figure 7.14), where the θ-dynamics oscillates between θ_1 and θ_2. The turning points u_1 and u_2 are roots of the function

$$F(u) = (1-u^2)\,[\epsilon - W(u)] \;=\; u^3 \;-\; (\epsilon+\beta^2)\,u^2 \;-\; (1-2\,\alpha\beta)\,u \;+\; (\epsilon-\alpha^2). \qquad (7.61)$$

Although a third root u_3 exists for $F(u) = 0$, it is unphysical since $u_3 > 1$.

Since the azimuthal frequencies at the turning points are expressed as

$$\frac{d\varphi_1}{d\tau} \;=\; \frac{\alpha-\beta\,u_1}{1-u_1^2} \quad\text{and}\quad \frac{d\varphi_2}{d\tau} \;=\; \frac{\alpha-\beta\,u_2}{1-u_2^2},$$

where $\alpha-\beta\,u_1 > \alpha-\beta\,u_2$, we can study the three cases for nutation numerically investigated below Eqs. (7.52); here, we assume that both $\alpha = b\,\beta$ and β are positive. In Case I ($\alpha > \beta\,u_2$), the precession frequency $d\varphi/d\tau$ is strictly positive for $u_1 \le u \le u_2$ and nutation proceeds monotonically. In Case II ($\alpha = \beta\,u_2$), the precession frequency $d\varphi/d\tau$ is positive for $u_1 \le u < u_2$ and vanishes at $u = u_2$; nutation in this Case exhibits a cusp at θ_2. In Case III ($\alpha < \beta\,u_2$), the precession frequency $d\varphi/d\tau$ reverses its sign at $u_r = \alpha/\beta = b$ or $\theta_2 < \theta_r = \arccos(b) < \theta_1$.

7.3.5 Stability of the Sleeping Top

Let us consider the case where a symmetric top with one fixed point is launched with initial conditions $\theta_0 \neq 0$ and $\dot{\theta}_0 = \dot{\varphi}_0 = 0$, with $\dot{\psi}_0 \neq 0$. In this case, the invariant canonical momenta are

$$p_\psi \;=\; I_3\,\dot{\psi}_0 \quad \text{and} \quad p_\varphi \;=\; p_\psi\,\cos\theta_0.$$

These initial conditions ($u_0 = \alpha/\beta, \dot{u}_0 = 0$), therefore, imply from Eq. (7.59) that $\epsilon = u_0$ and that the energy equation (7.59) now becomes

$$\left(\frac{du}{d\tau}\right)^2 \;=\; \Big[\,(1-u^2)\;-\;\beta^2\,(u_0-u)\,\Big]\;(u_0-u). \tag{7.62}$$

Next, we consider the case of the *sleeping* top for which an additional initial condition is $\theta_0 = 0$ (and $u_0 = 1$). Thus Eq. (7.62) becomes

$$\left(\frac{du}{d\tau}\right)^2 \;=\; \Big(1\;+\;u\;-\;\beta^2\Big)\;(1-u)^2. \tag{7.63}$$

The sleeping top has the following equilibrium points (where $\dot{u} = 0$): $u_1 = 1$ and $u_2 = \beta^2 - 1$. We now investigate the stability of the equilibrium point $u_1 = 1$ by writing $u = 1 - \delta$ (with $\delta \ll 1$) so that Eq. (7.63) becomes

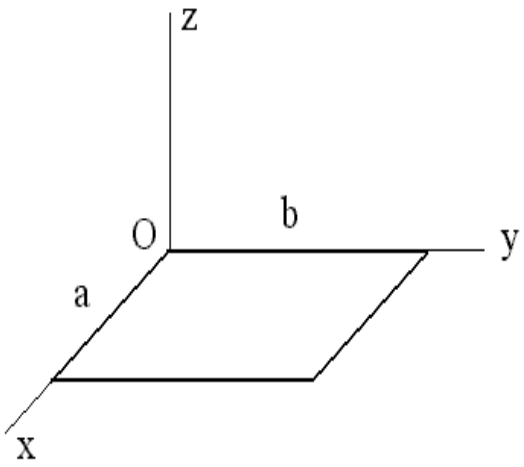
$$\frac{d\delta}{d\tau} \;=\; \Big(2-\beta^2\Big)^{\frac{1}{2}}\;\delta.$$

The solution of this equation is exponential (and, therefore, u_1 is unstable) if $\beta^2 < 2$ or oscillatory (and, therefore, u_1 is stable) if $\beta^2 > 2$. Note that in the latter case, the condition $\beta^2 > 2$ implies that the second equilibrium point $u_2 = \beta^2 - 1 > 1$ is unphysical. We, therefore, see that stability of the sleeping top requires a large spinning frequency ω_3; in the presence of friction, the spinning frequency slows down and ultimately the sleeping top becomes unstable.

7.4 Problems

Problem 1

Consider a thin homogeneous rectangular plate of mass M and area ab that lies on the (x, y)-plane.



(a) Show that the inertia tensor (calculated in the reference frame with its origin at point O in the Figure above) takes the form

$$\mathbf{I} = \begin{pmatrix} A & -C & 0 \\ -C & B & 0 \\ 0 & 0 & A+B \end{pmatrix},$$

and find suitable expressions for A, B, and C in terms of M, a, and b.

(b) Show that by performing a rotation of the coordinate axes about the z-axis through an angle θ, the new inertia tensor is

$$\mathbf{I}'(\theta) = \mathsf{R}(\theta) \cdot \mathbf{I} \cdot \mathsf{R}^T(\theta) = \begin{pmatrix} A' & -C' & 0 \\ -C' & B' & 0 \\ 0 & 0 & A'+B' \end{pmatrix},$$

where

$$\begin{aligned} A' &= A\cos^2\theta + B\sin^2\theta - C\sin 2\theta \\ B' &= A\sin^2\theta + B\cos^2\theta + C\sin 2\theta \\ C' &= C\cos 2\theta - \frac{1}{2}(B-A)\sin 2\theta. \end{aligned}$$

(c) When

$$\theta \;=\; \frac{1}{2}\,\arctan\left(\frac{2C}{B-A}\right),$$

the off-diagonal component C' vanishes and the $x'-$ and $y'-$axes become principal axes. Calculate expressions for A' and B' in terms of M, a, and b for this particular angle.

(d) Calculate the inertia tensor $\mathbf{I}_{CM}$ in the CM frame by using the Parallel-Axis Theorem and show that

$$I^x_{CM} \;=\; \frac{M\,b^2}{12},\;\; I^y_{CM} \;=\; \frac{M\,a^2}{12},\;\; \text{and}\;\; I^z_{CM} \;=\; \frac{M}{12}\left(b^2 \;+\; a^2\right).$$

Problem 2

(a) The Euler equation for an asymmetric top ($I_1 > I_2 > I_3$) with $L^2 = 2\,I_2 K$ is

$$\dot{\omega}_2 \;=\; \alpha\,\left(\Omega^2 \;-\; \omega_2^2\right),$$

where

$$\Omega^2 \;=\; \frac{2K}{I_2} \quad \text{and} \quad \alpha \;=\; \sqrt{\left(1 \;-\; \frac{I_2}{I_1}\right)\left(\frac{I_2}{I_3} \;-\; 1\right)}$$

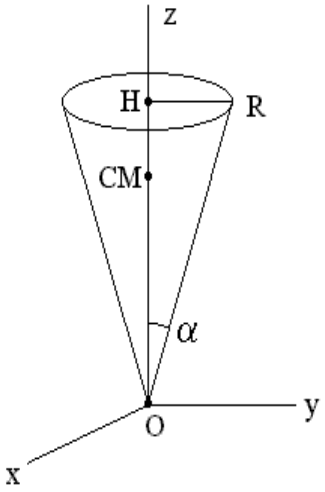
Solve for $\omega_2(t)$ with the initial condition $\omega_2(0) = 0$.

(b) Use the solution $\omega_2(t)$ found in Part (a) to find the solutions $\omega_1(t)$ and $\omega_3(t)$ given by Eqs. (7.35).

Problem 3

(a) Consider a circular cone of height H and base radius $R = H\,\tan\alpha$ with uniform mass

density $\rho = 3\,M/(\pi\,HR^2)$.



Show that the non-vanishing components of the inertia tensor **I** calculated from the vertex O of the cone are

$$I_{xx} \;=\; I_{yy} \;=\; \frac{3}{5}\,M\left(H^2 \;+\; \frac{R^2}{4}\right) \quad \text{and} \quad I_{zz} \;=\; \frac{3}{10}\,M\,R^2$$

(b) Show that the principal moments of inertia calculated in the CM frame (located at a height $h = 3H/4$ on the symmetry axis) are

$$I_1 \;=\; I_2 \;=\; \frac{3}{20}\,M\left(R^2 \;+\; \frac{H^2}{4}\right) \quad \text{and} \quad I_3 \;=\; \frac{3}{10}\,M\,R^2$$

Problem 4

Show that the Euler basis vectors $(\hat{\mathsf{e}}_1, \hat{\mathsf{e}}_2, \hat{\mathsf{e}}_3)$ are expressed as shown in Eq. (7.41).

Chapter 8

Normal-Mode Analysis

8.1 Stability of Equilibrium Points

A nonlinear force equation $m\,\ddot{x} = f(x)$ has equilibrium points (labeled x_0) when $f(x_0)$ vanishes. The stability of the equilibrium point x_0 is determined by the sign of $f'(x_0)$: the equilibrium point x_0 is stable if $f'(x_0) < 0$ or it is unstable if $f'(x_0) > 0$. Since $f(x)$ is also derived from a potential $V(x)$ as $f(x) = -\,V'(x)$, we say that the equilibrium point x_0 is stable (or unstable) if $V''(x_0)$ is positive (or negative).

8.1.1 Bead on a Rotating Hoop

In Chap. 2, we considered the problem of a bead of mass m sliding freely on a hoop of radius r rotating with angular velocity ω_0 in a constant gravitational field with acceleration g. The Lagrangian for this system is

$$L(\theta,\dot{\theta}) \;=\; \frac{m}{2}\,r^2\dot{\theta}^2 \;+\; \left(\frac{m}{2}\,r^2\omega_0^2\,\sin^2\theta \;+\; mgr\,\cos\theta\right) \;=\; \frac{m}{2}\,r^2\dot{\theta}^2 \;-\; V(\theta),$$

where $V(\theta)$ denotes the effective potential, and the Euler-Lagrange equation for θ is

$$mr^2\,\ddot{\theta} \;=\; -\,V'(\theta) \;=\; -\,mr^2\omega_0^2\,\sin\theta\,(\nu-\cos\theta), \tag{8.1}$$

where $\nu = g/(r\,\omega_0^2)$. The equilibrium points of Eq. (8.1) are $\theta = 0$ (for all values of ν) and $\theta = \arccos(\nu)$ if $\nu < 1$. The stability of the equilibrium point $\theta = \theta_0$ is determined by the sign of

$$V''(\theta_0) \;=\; mr^2\omega_0^2\,\left[\nu\,\cos\theta_0 \;-\; \left(2\,\cos^2\theta_0 - 1\right)\right].$$

Hence,

$$V''(0) \;=\; mr^2\omega_0^2\,(\nu \;-\; 1) \tag{8.2}$$

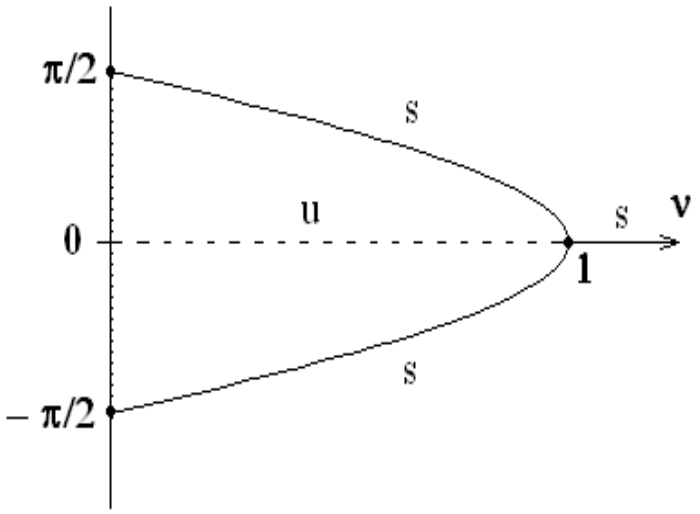


Figure 8.1: Bifurcation tree for the bead on a rotating-hoop problem

is positive (i.e., $\theta = 0$ is stable) if $\nu > 1$ or negative (i.e., $\theta = 0$ is unstable) if $\nu < 1$. In the latter case, when $\nu < 1$ and the second equilibrium point $\theta_0 = \arccos(\nu)$ is possible, we find

$$V''(\theta_0) \;=\; mr^2\omega_0^2 \left[\nu^2 \;-\; \left(2\,\nu^2 - 1\right)\right] \;=\; mr^2\omega_0^2 \left(1 - \nu^2\right) \;>\; 0, \tag{8.3}$$

and thus the equilibrium point $\theta_0 = \arccos(\nu)$ is stable when $\nu < 1$. The stability of the bead on a rotating hoop is displayed on the *bifurcation* diagram (see Figure 7.14) which shows the stable regime bifurcates at $\nu = 1$.

8.1.2 Circular Orbits in Central-Force Fields

The radial force equation

$$\mu\,\ddot{r} \;=\; \frac{\ell^2}{\mu\,r^3} \;-\; \frac{k}{r^n} \;=\; -\,V'(r),$$

studied in Chap. 4 for a central-force field $F(r) = -\,k\,r^{-n}$ (here, μ is the reduced mass of the system, the azimuthal angular momentum ℓ is a constant of the motion, and k is a constant), has the equilibrium point at $r = \rho$ defined by the relation

$$\rho^{n-3} \;=\; \frac{\ell^2}{\mu\,k}. \tag{8.4}$$

The second derivative of the effective potential is

$$V''(r) \;=\; \frac{\ell^2}{\mu\,r^4}\left(3 \;-\; n\,\frac{k\,\mu}{\ell^2\,r^{n-3}}\right),$$

which becomes

$$V''(\rho) \;=\; \frac{\ell^2}{\mu\,\rho^4}\,(3-n). \tag{8.5}$$

Hence, $V''(\rho)$ is positive if $n < 3$, and thus circular orbits are stable in central-force fields $F(r) = -\,k\,r^{-n}$ if $n < 3$.

8.2 Small Oscillations about Stable Equilibria

Once an equilibrium point x_0 is shown to be stable, i.e., $f'(x_0) < 0$ or $V''(x_0) > 0$, we may expand $x = x_0 + \delta x$ about the equilibrium point (with $\delta x \ll x_0$) to find the *linearized* force equation

$$m\,\delta\ddot{x} \;=\; -\,V''(x_0)\,\delta x, \tag{8.6}$$

which has oscillatory behavior with frequency

$$\omega(x_0) \;=\; \sqrt{\frac{V''(x_0)}{m}}.$$

We first look at the problem of a bead on a rotating hoop, where the frequency of small oscillations $\omega(\theta_0)$ is either given in Eq. (8.2) as

$$\omega(0) \;=\; \sqrt{\frac{V''(0)}{mr^2}} \;=\; \omega_0\,\sqrt{\nu-1}$$

for $\theta_0 = 0$ and $\nu > 1$, or is given in Eq. (8.3) as

$$\omega(\theta_0) \;=\; \sqrt{\frac{V''(\theta_0)}{mr^2}} \;=\; \omega_0\,\sqrt{1-\nu^2}$$

for $\theta_0 = \arccos(\nu)$ and $\nu < 1$.

Next, we look at the frequency of small oscillations about the stable circular orbit in a central-force field $F(r) = -\,k\,r^{-n}$ (with $n < 3$). Here, from Eq. (8.5), we find

$$\omega \;=\; \sqrt{\frac{V''(\rho)}{\mu}} \;=\; \sqrt{\frac{k\,(3-n)}{\mu\,\rho^{n+1}}},$$

where $\ell^2 = \mu\,k\,\rho^{3-n}$ was used. We note that for the Kepler problem ($n = 2$), the period of small oscillations $T = 2\pi/\omega$ is expressed as

$$T^2 \;=\; \frac{(2\pi)^2\,\mu}{k}\,\rho^3,$$

which is precisely the statement of Kepler's Third Law for circular orbits. Hence, a small perturbation of a stable Keplerian circular orbit does not change its orbital period.

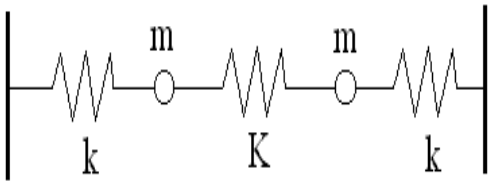


Figure 8.2: Coupled masses and springs

8.3 Coupled Oscillations and Normal-Mode Analysis

8.3.1 Coupled Simple Harmonic Oscillators

We begin our study of linearily-coupled oscillators by considering the following coupled system comprised of two block-and-spring systems (with identical mass m and identical spring constant k) coupled by means of a spring of constant K. The coupled equations are

$$m\,\ddot{x}_1 \;=\; -\,(k+K)\,x_1 \;+\; K\,x_2 \quad \text{and} \quad m\,\ddot{x}_2 \;=\; -\,(k+K)\,x_2 \;+\; K\,x_1. \tag{8.7}$$

The solutions for $x_1(t)$ and $x_2(t)$ by following a method known as the normal-mode analysis. First, we write $x_1(t)$ and $x_2(t)$ in the normal-mode representation

$$x_1(t) \;=\; \overline{x}_1\,e^{-\,i\omega t} \quad \text{and} \quad x_2(t) \;=\; \overline{x}_2\,e^{-\,i\omega t}, \tag{8.8}$$

where $\overline{x}_1$ and $\overline{x}_2$ are constants and the eigenfrequency ω is to be solved in terms of the system parameters (m, k, K). Next, substituting the normal-mode representation into Eq. (8.7), we obtain the following normal-mode matrix equation

$$\begin{pmatrix} \omega^2\,m-(k+K) & K \\ K & \omega^2\,m-(k+K) \end{pmatrix} \begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \end{pmatrix} \;=\; 0. \tag{8.9}$$

To obtain a non-trivial solution $\overline{x}_1 \neq 0 \neq \overline{x}_2$, the determinant of the matrix in Eq. (8.9) is required to vanish, which yields the characteristic polynomial

$$[\omega^2\,m-(k+K)]^2 \;-\; K^2 \;=\; 0,$$

whose solutions are the eigenfrequencies

$$\omega_{\pm}^2 \;=\; \frac{(k+K)}{m} \;\pm\; \frac{K}{m}.$$

If we insert $\omega_+^2 = (k+2K)/m$ into the matrix equation (8.9), we find

$$\begin{pmatrix} K & K \\ K & K \end{pmatrix} \begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \end{pmatrix} \;=\; 0,$$

which implies that $\overline{x}_2 = -\,\overline{x}_1$, and thus the *eigenfrequency* ω_+ is associated with an *anti-symmetric* coupled motion. If we insert $\omega_-^2 \;=\; k/m$ into the matrix equation (8.9), we find

$$\begin{pmatrix} -\,K & K \\ K & -\,K \end{pmatrix} \begin{pmatrix} \overline{x}_1 \\ \overline{x}_2 \end{pmatrix} \;=\; 0,$$

which implies that $\overline{x}_2 = \overline{x}_1$, and thus the eigenfrequency ω_- is associated with a *symmetric* coupled motion.

Lastly, we construct the normal coordinates η_+ and η_-, which satisfy the condition $\ddot{\eta}_\pm = -\,\omega_\pm^2\,\eta_\pm$. From the discussion above, we find

$$\eta_-(t) \;=\; x_1(t) + x_2(t) \quad \text{and} \quad \eta_+(t) \;=\; x_1(t) - x_2(t). \tag{8.10}$$

The solutions $\eta_\pm(t)$ are of the form

$$\eta_\pm \;=\; A_\pm\,\cos(\omega_\pm t + \varphi_\pm),$$

where $A_\pm$ and $\varphi_\pm$ are constants (determined from initial conditions). The general solution of Eqs. (8.7) can, therefore, be written explicitly in terms of the normal coordinates $\eta_\pm$ as follows

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} \;=\; \frac{A_-}{2}\,\cos(\omega_- t + \varphi_-) \;\pm\; \frac{A_+}{2}\,\cos(\omega_+ t + \varphi_+).$$

8.3.2 Nonlinear Coupled Oscillators

We now consider the following system composed of two pendula of identical length ℓ but different masses m_1 and m_2 coupled by means of a spring of constant k in the presence of a gravitational field of constant acceleration g. (The distance D between the two points of attach of the pendula is equal to the length of the spring in its relaxed state and we assume, for simplicity, that the masses always stay on the same horizontal line.) Using the generalized coordinates (θ_1, θ_2) defined in the Figure above, the Lagrangian for this system is

$$\begin{aligned} L \;&=\; \frac{\ell^2}{2}\,\left(m_1\,\dot{\theta}_1^2 + m_2\,\dot{\theta}_2^2\right) \;-\; g\ell\,\left\{\, m_1\,(1-\cos\theta_1) + m_2\,(1-\cos\theta_2)\,\right\} \\ &\quad -\; \frac{k\ell^2}{2}\,(\sin\theta_1 - \sin\theta_2)^2\,, \end{aligned}$$

and the nonlinear coupled equations of motion are

$$\begin{aligned} m_1\,\ddot{\theta}_1 \;&=\; -\,m_1\,\omega_g^2\,\sin\theta_1 \;-\; k\,(\sin\theta_1 - \sin\theta_2)\,\cos\theta_1, \\ m_2\,\ddot{\theta}_2 \;&=\; -\,m_2\,\omega_g^2\,\sin\theta_2 \;+\; k\,(\sin\theta_1 - \sin\theta_2)\,\cos\theta_2, \end{aligned}$$

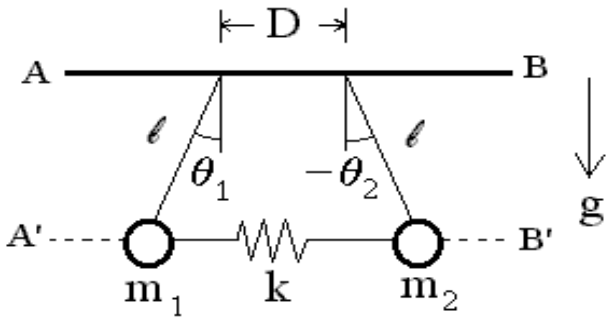


Figure 8.3: Coupled pendula

where $\omega_g^2 = g/\ell$.

It is quite clear that the equilibrium point is $\theta_1 = 0 = \theta_2$ and expansion of the coupled equations about this equilibrium yields the coupled linear equations

$$\begin{aligned} m_1\,\ddot{q}_1 &= -\,m_1\,\omega_g^2\,q_1 \;-\; k\,(q_1 - q_2), \\ m_2\,\ddot{q}_2 &= -\,m_2\,\omega_g^2\,q_2 \;+\; k\,(q_1 - q_2), \end{aligned}$$

where $\theta_1 = q_1 \ll 1$ and $\theta_2 = q_2 \ll 1$. The normal-mode matrix associated with these coupled linear equations is

$$\begin{pmatrix} (\omega^2 - \omega_g^2)\,m_1 - k & k \\ k & (\omega^2 - \omega_g^2)\,m_2 - k \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} \;=\; 0,$$

and the characteristic polynomial is

$$M\,\left[\,(\omega^2 - \omega_g^2)\,\mu \;-\; k\,\right]\,(\omega^2 - \omega_g^2) \;=\; 0,$$

where $\mu = m_1 m_2/M$ is the reduced mass for the system and $M = m_1 + m_2$ is the total mass. The eigenfrequencies are thus

$$\omega_-^2 \;=\; \omega_g^2 \quad \text{and} \quad \omega_+^2 \;=\; \omega_g^2 \;+\; \frac{k}{\mu}.$$

The normal coordinates $\eta_\pm$ are expressed in terms of (q_1, q_2) as $\eta_\pm = a_\pm\,q_1 + b_\pm\,q_2$, where $a_\pm$ and $b_\pm$ are constant coefficients determined from the condition $\ddot{\eta}_\pm = -\,\omega_\pm^2\,\eta_\pm$. From

this condition we find

$$\left(\frac{k}{m_1}+\omega_g^2\right) \;-\; \frac{b_\pm}{a_\pm}\,\frac{k}{m_2} \;=\; \omega_\pm^2 \;=\; \left(\frac{k}{m_2}+\omega_g^2\right) \;-\; \frac{a_\pm}{b_\pm}\,\frac{k}{m_1}.$$

For the eigenfrequency $\omega_-^2 = \omega_g^2$, we find $b_-/a_- = m_2/m_1$, and thus we may choose

$$\eta_- \;=\; \frac{m_1}{M}\,q_1 \;+\; \frac{m_2}{M}\,q_2,$$

which represents the center of mass position for the system. For the eigenfrequency $\omega_+^2 = \omega_g + k/\mu$, we find $b_+/a_+ = -\,1$, and thus we may choose

$$\eta_+ \;=\; q_1 \;-\; q_2.$$

Lastly, we may solve for q_1 and q_2 as

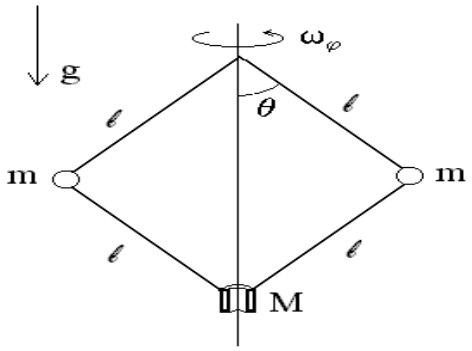
$$q_1 \;=\; \eta_- \;+\; \frac{m_2}{M}\,\eta_+ \quad\text{and}\quad q_2 \;=\; \eta_- \;-\; \frac{m_1}{M}\,\eta_+,$$

where $\eta_\pm = A_\pm\,\cos(\omega_\pm\,t + \varphi_\pm)$ are general solutions of the normal-mode equations $\ddot{\eta}_\pm = -\,\omega_\pm^2\,\eta_\pm$.

8.4 Problems

Problem 1

The following compound pendulum is composed of two identical masses m attached by massless rods of identical length ℓ to a ring of mass M, which is allowed to slide up and down along a vertical axis in a gravitational field with constant g. The entire system rotates about the vertical axis with an azimuthal angular frequency ω_φ.



(a) Show that the Lagrangian for the system can be written as

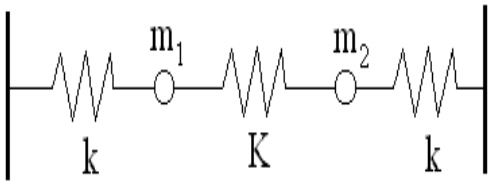
$$L(\theta, \dot{\theta}) \;=\; \ell^2 \dot{\theta}^2 \left(m + 2M\, \sin^2\theta\right) \;+\; m\ell^2\, \omega_\varphi^2\, \sin^2\theta \;+\; 2\,(m+M) g\ell\, \cos\theta$$

(b) Identify the equilibrium points for the system and investigate their stability.

(c) Determine the frequency of small oscillations about each stable equilibrium point found in Part (b).

Problem 2

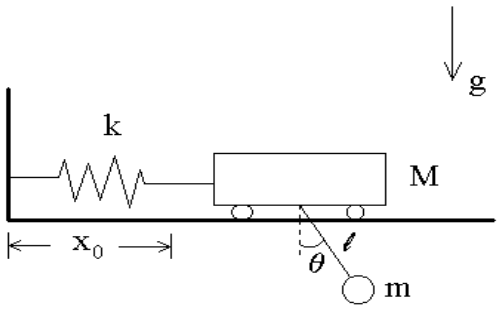
Consider the same problem as in Sec. (8.3.1) but now with different masses ($m_1 \neq m_2$).



Calculate the eigenfrequencies and eigenvectors (normal coordinates) for this system.

Problem 3

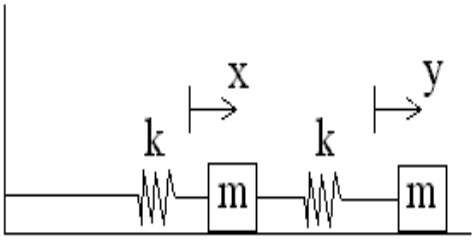
Find the eigenfrequencies associated with small oscillations of the system shown below.



Problem 4

Two blocks of identical mass m are attached by massless springs (with identical spring

constant k) as shown in the Figure below.



The Lagrangian for this system is

$$L(x, \dot{x};\, y, \dot{y}) \;=\; \frac{m}{2}\left(\dot{x}^2 \;+\; \dot{y}^2\right) \;-\; \frac{k}{2}\left[\, x^2 \;+\; (y-x)^2 \,\right],$$

where x and y denote departures from equilibrium.

(a) Derive the Euler-Lagrange equations for x and y.

(b) Show that the eigenfrequencies for small oscillations for this system are

$$\omega_{\pm}^2 \;=\; \frac{\omega_k^2}{2}\,\left(\,3 \;\pm\; \sqrt{5}\,\right),$$

where $\omega_k^2 = k/m$.

(c) Show that the eigenvectors associated with the eigenfrequencies $\omega_{\pm}$ are represented by the relations

$$\overline{y}_{\pm} \;=\; \frac{1}{2}\left(\,1 \;\mp\; \sqrt{5}\,\right)\overline{x}_{\pm}$$

where $(\overline{x}_{\pm},\ \overline{y}_{\pm})$ represent the normal-mode amplitudes.

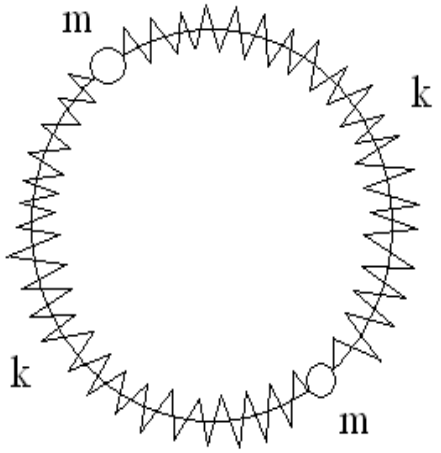
Problem 5

An infinite sheet with surface mass density σ has a hole of radius R cut into it. A particle of mass m sits (in equilibrium) at the center of the circle. Assuming that the

sheet lies on the (x, y)-plane (with the hole centered at the origin) and that the particle is displaced by a small amount $z \ll R$ along the z-axis, calculate the frequency of small oscillations.

Problem 6

Two identical masses are connected by two identical massless springs and are constrained to move on a circle (see Figure below). Of course, the two masses are in equilibrium when they are diametrically opposite points on the circle.



Solve for the normal modes of the system.

Chapter 9

Continuous Lagrangian Systems

9.1 Waves on a Stretched String

9.1.1 Wave Equation

The equation describing waves propagating on a stretched string of constant linear mass density ρ under constant tension T is

$$\rho \, \frac{\partial^2 u(x,t)}{\partial t^2} \;=\; T \, \frac{\partial^2 u(x,t)}{\partial x^2}, \tag{9.1}$$

where $u(x,t)$ denotes the amplitude of the wave at position x along the string at time t. General solutions to this equation involve arbitrary functions $g(x \pm v\,t)$, where $v = \sqrt{T/\rho}$ represents the speed of waves propagating on the string. Indeed, we find

$$\rho \, \partial_t^2 g(x \pm v\,t) \;=\; \rho \, v^2 \; g' \;=\; T \; g' \;=\; T \, \partial_x^2 g(x \pm v\,t).$$

The interpretation of the two different signs is that $g(x - v\,t)$ represents a wave propagating to the right while $g(x + v\,t)$ represents a wave propagating to the left. The general solution of the wave equation (9.1) is

$$u(x,t) \;=\; A_- \; g(x - v\,t) \;+\; A_+ \; g(x + v\,t),$$

where $A_\pm$ are arbitrary constants.

9.1.2 Lagrangian Formalism

The question we now ask is whether the wave equation (9.1) can be derived from a variational principle

$$\delta \int \; \mathcal{L}(u, \partial_t u, \partial_x u;\; x, t) \; dx \, dt \;=\; 0, \tag{9.2}$$

where the Lagrangian *density* $\mathcal{L}(u, \partial_t u, \partial_x u;\ x, t)$ is a function of the dynamical variable $u(x,t)$ and its space-time derivatives. Here, the variation of the Lagrangian density $\mathcal{L}$ is expressed as

$$\delta\mathcal{L} \ = \ \delta u \, \frac{\partial \mathcal{L}}{\partial u} \ + \ \partial_t \delta u \, \frac{\partial \mathcal{L}}{\partial(\partial_t u)} \ + \ \partial_x \delta u \, \frac{\partial \mathcal{L}}{\partial(\partial_x u)},$$

where $\delta u(x,t)$ is a general variation of $u(x,t)$ subject to the condition that it vanishes at the integration boundaries in Eq. (9.2). By re-arranging terms, the variation of $\mathcal{L}$ can be written as

$$\begin{aligned} \delta\mathcal{L} \ &= \ \delta u \left\{ \frac{\partial \mathcal{L}}{\partial u} \ - \ \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial(\partial_t u)} \right) \ - \ \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial(\partial_x u)} \right) \right\} \\ &\quad + \ \frac{\partial}{\partial t}\left(\delta u \, \frac{\partial \mathcal{L}}{\partial(\partial_t u)} \right) \ + \ \frac{\partial}{\partial x}\left(\delta u \, \frac{\partial \mathcal{L}}{\partial(\partial_x u)} \right). \end{aligned} \tag{9.3}$$

When we insert this expression for $\delta\mathcal{L}$ into the variational principle (9.2), we obtain

$$\int dx\, dt \ \delta u \left\{ \frac{\partial \mathcal{L}}{\partial u} \ - \ \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial(\partial_t u)} \right) \ - \ \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial(\partial_x u)} \right) \right\} \ = \ 0, \tag{9.4}$$

where the last two terms in Eq. (9.3) cancel out because δu vanishes on the integration boundaries. Since the variational principle (9.4) is true for general variations δu, we obtain the Euler-Lagrange equation for the dynamical field $u(x,t)$:

$$\frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial(\partial_t u)} \right) \ + \ \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial(\partial_x u)} \right) \ = \ \frac{\partial \mathcal{L}}{\partial u}. \tag{9.5}$$

9.1.3 Lagrangian Description for Waves on a Stretched String

The question we posed earlier now focuses on deciding what form the Lagrangian density must take. Here, the answer is surprisingly simple: the kinetic energy density of the wave is $\rho\ (\partial_t u)^2/2$, while the potential energy density is $T\ (\partial_x u)^2/2$, and thus the Lagrangian density for waves on a stretched string is

$$\mathcal{L}(u, \partial_t u, \partial_x u;\ x, t) \ = \ \frac{\rho}{2}\left(\frac{\partial u}{\partial t} \right)^2 \ - \ \frac{T}{2}\left(\frac{\partial u}{\partial x} \right)^2. \tag{9.6}$$

Since $\partial\mathcal{L}/\partial u = 0$, we find

$$\begin{aligned} \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}}{\partial(\partial_t u)} \right) \ &= \ \frac{\partial}{\partial t}\left(\rho \, \frac{\partial u}{\partial t} \right) \ = \ \rho \, \frac{\partial^2 u}{\partial t^2}, \\ \frac{\partial}{\partial x}\left(\frac{\partial \mathcal{L}}{\partial(\partial_x u)} \right) \ &= \ \frac{\partial}{\partial x}\left(-\, T \, \frac{\partial u}{\partial x} \right) \ = \ -\, T \, \frac{\partial^2 u}{\partial x^2}, \end{aligned}$$

and Eq. (9.1) is indeed represented as an Euler-Lagrange equation (9.5) in terms of the Lagrangian density (9.6).

9.2 General Variational Principle for Field Theory

The simple example of waves on a stretched string allows us to view the Euler-Lagrange equation (9.5) as a generalization of the Euler-Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) \;=\; \frac{\partial L}{\partial q^i},$$

in terms of the generalized coordinates q^i. We now spend some time investigating the Lagrangian description of continuous systems, in which the dynamical variables are fields $\psi(\mathbf{x}, t)$ instead of spatial coordinates $\mathbf{x}$.

9.2.1 Action Functional

Classical and quantum field theories rely on variational principles based on the existence of action functionals. The typical action functional is of the form

$$\mathcal{A}[\psi] \;=\; \int d^4x\; \mathcal{L}(\psi,\; \partial_\mu \psi), \tag{9.7}$$

where the wave function $\psi(\mathbf{x}, t)$ represents the state of the system at position $\mathbf{x}$ and time t while the entire physical content of the theory is carried by the Lagrangian density $\mathcal{L}$.

The variational principle is based on the stationarity of the action functional

$$\delta\mathcal{A}[\psi] \;=\; \mathcal{A}[\psi + \delta\psi] \;-\; \mathcal{A}[\psi] \;=\; \int\; \delta\mathcal{L}(\psi,\; \partial_\mu\psi)\; d^4x,$$

where $\partial_\mu = (c^{-1}\partial_t,\; \nabla)$ and the metric tensor is defined as $g^{\mu\nu} = \mathrm{diag}(-1, +1, +1, +1)$.[1] Here, the variation of the Lagrangian density is

$$\begin{aligned} \delta\mathcal{L} \;&=\; \frac{\partial\mathcal{L}}{\partial\psi}\,\delta\psi \;+\; \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\,\partial_\mu\delta\psi \\ &\equiv\; \delta\psi\left[\;\frac{\partial\mathcal{L}}{\partial\psi} \;-\; \frac{\partial}{\partial x^\mu}\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\right)\;\right] \;+\; \frac{\partial\Lambda^\mu}{\partial x^\mu}, \end{aligned} \tag{9.8}$$

where

$$\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\,\partial_\mu\delta\psi \;=\; \frac{\partial\mathcal{L}}{\partial(\nabla\psi)}\cdot\nabla\delta\psi \;+\; \frac{\partial\mathcal{L}}{\partial(\partial_t\psi)}\,\partial_t\delta\psi,$$

and the exact space-time divergence $\partial_\mu\Lambda^\mu$ is obtained by rearranging terms, with

$$\Lambda^\mu \;=\; \delta\psi\,\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)} \quad\text{and}\quad \frac{\partial\Lambda^\mu}{\partial x^\mu} \;=\; \frac{\partial}{\partial t}\left(\delta\psi\,\frac{\partial\mathcal{L}}{\partial(\partial_t\psi)}\right) \;+\; \nabla\cdot\left(\delta\psi\,\frac{\partial\mathcal{L}}{\partial(\nabla\psi)}\right).$$

[1] For two four-vectors $A^\mu = (A^0, \mathbf{A})$ and $B^\mu = (B^0, \mathbf{B})$, we have $A\cdot B = A_\mu\,B^\mu = \mathbf{A}\cdot\mathbf{B} - A^0\,B^0$, where $A_0 = -\,A^0$.

The variational principle $\delta\mathcal{A}[\psi] = 0$ then yields

$$0 \;=\; \int d^4x\, \delta\psi \left[\frac{\partial\mathcal{L}}{\partial\psi} \;-\; \frac{\partial}{\partial x^\mu}\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\right)\right],$$

where the exact divergence $\partial_\mu\Lambda^\mu$ drops out under the assumption that the variation $\delta\psi$ vanish on the integration boundaries. Following the standard rules of Calculus of Variations, the Euler-Lagrange equation for the wave function ψ is

$$\frac{\partial}{\partial x^\mu}\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\right) \;=\; \frac{\partial}{\partial t}\left(\frac{\partial\mathcal{L}}{\partial(\partial_t\psi)}\right) \;+\; \nabla\cdot\left(\frac{\partial\mathcal{L}}{\partial(\nabla\psi)}\right) \;=\; \frac{\partial\mathcal{L}}{\partial\psi}. \tag{9.9}$$

9.2.2 Noether Method and Conservation Laws

Since the Euler-Lagrange equations (9.9) hold true for arbitrary field variations $\delta\psi$, the variation of the Lagrangian density $\mathcal{L}$ is now expressed as the Noether equation

$$\delta\mathcal{L} \;\equiv\; \partial_\mu\Lambda^\mu \;=\; \frac{\partial}{\partial x^\mu}\left[\,\delta\psi\,\frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\,\right]. \tag{9.10}$$

which associates symmetries with conservation laws $\partial_\mu\,\mathcal{J}^\mu_a = 0$, where the index a denotes the possibility of conserved four-vector quantities (see below).

Energy-Momentum Conservation Law

The conservation of energy-momentum (a four-vector quantity!) involves symmetry of the Lagrangian with respect to constant *space-time* translations ($x^\nu \to x^\nu + \delta x^\nu$). Here, the variation $\delta\psi$ is no longer arbitrary but is required to be of the form

$$\delta\psi \;=\; -\,\delta x^\nu\,\partial_\nu\psi \tag{9.11}$$

while the variation $\delta\mathcal{L}$ is of the form

$$\delta\mathcal{L} \;=\; -\,\delta x^\nu\left[\,\partial_\nu\mathcal{L} \;-\; (\partial_\nu\mathcal{L})_\psi\,\right], \tag{9.12}$$

where $(\partial_\nu\mathcal{L})_\psi$ denotes a explicit space-time derivative of $\mathcal{L}$ at constant ψ. The Noether equation (9.10) can now be written as

$$\frac{\partial}{\partial x^\mu}\left(\mathcal{L}\;g^\mu{}_\nu \;-\; \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\,\partial_\nu\psi\right) \;=\; \left(\frac{\partial\mathcal{L}}{\partial x^\nu}\right)_\psi.$$

If the Lagrangian is explicitly independent of the space-time coordinates, i.e., $(\partial_\nu\mathcal{L})_\psi = 0$, the energy-momentum conservation law $\partial_\mu\,T^\mu{}_\nu = 0$ is written in terms of the energy-momentum tensor

$$T^\mu{}_\nu \;\equiv\; \mathcal{L}\;g^\mu{}_\nu \;-\; \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\,\partial_\nu\psi. \tag{9.13}$$

We note that the derivation of the energy-momentum conservation law is the same for classical and quantum fields. A similar procedure would lead to the conservation of angular momentum but this derivation is beyond the scope of the present Notes and we move on instead to an important conservation in wave dynamics.

Wave-Action Conservation Law

Waves are known to exist on a great variety of media. When waves are supported by a spatially nonuniform or time-dependent medium, the conservation law of energy or momentum no longer apply and instead energy or momentum is transfered between the medium and the waves. There is however one conservation law which still applies and the quantity being conserved is known as *wave action*.

The derivation of a wave-action conservation law differs for classical fields and quantum fields. The difference is related to the fact that whereas classical fields are generally represented by real-valued wave functions (i.e., $\psi^* = \psi$), the wave functions of quantum field theories are generally complex-valued (i.e., $\psi^* \neq \psi$).

The first step in deriving a wave-action conservation law in classical field theory involves transforming the real-valued wave function ψ into a complex-valued wave function ψ. Next, variations of ψ and its complex conjugate ψ^* are of the form

$$\delta\psi \;\equiv\; i\epsilon\,\psi \;\text{ and }\; \delta\psi^* \;\equiv\; -\,i\epsilon\,\psi^*, \tag{9.14}$$

Lastly, we transform the classical Lagrangian density $\mathcal{L}$ into a real-valued Lagrangian density $\mathcal{L}_R(\psi, \psi^*)$ such that $\delta\mathcal{L}_R \;\equiv\; 0$. The wave-action conservation law is, therefore, expressed in the form $\partial_\mu \mathcal{J}^\mu = 0$, where the wave-action four-density is

$$\mathcal{J}^\mu \;\equiv\; 2\,\mathrm{Im}\left[\,\psi\,\frac{\partial\mathcal{L}_R}{\partial(\partial_\mu\psi)}\,\right], \tag{9.15}$$

where $\mathrm{Im}[\cdots]$ denotes the imaginary part.

9.3 Variational Principle for Schroedinger Equation

A simple yet important example for a quantum field theory is provided by the Schroedinger equation for a spinless particle of mass m subjected to a real-valued potential energy function $V(\mathbf{x}, t)$. The Lagrangian density for the Schroedinger equation is given as

$$\mathcal{L}_R \;=\; -\,\frac{\hbar^2}{2m}\,|\nabla\psi|^2 \;+\; \frac{i\hbar}{2}\left(\psi^*\,\frac{\partial\psi}{\partial t} \;-\; \psi\,\frac{\partial\psi^*}{\partial t}\right) \;-\; V\,|\psi|^2. \tag{9.16}$$

The Schroedinger equation for ψ is derived as an Euler-Lagrange equation (9.9) in terms of ψ^*, where

$$\begin{aligned}
\frac{\partial \mathcal{L}_R}{\partial(\partial_t \psi^*)} &= -\frac{i\hbar}{2}\,\psi \;\rightarrow\; \frac{\partial}{\partial t}\left(\frac{\partial \mathcal{L}_R}{\partial(\partial_t \psi^*)}\right) = -\frac{i\hbar}{2}\,\frac{\partial \psi}{\partial t}, \\
\frac{\partial \mathcal{L}_R}{\partial(\nabla \psi^*)} &= -\frac{\hbar^2}{2m}\,\nabla\psi \;\rightarrow\; \nabla\cdot\left(\frac{\partial \mathcal{L}_R}{\partial(\nabla \psi^*)}\right) = -\frac{\hbar^2}{2m}\,\nabla^2\psi, \\
\frac{\partial \mathcal{L}_R}{\partial \psi^*} &= \frac{i\hbar}{2}\,\frac{\partial \psi}{\partial t} - V\,\psi,
\end{aligned}$$

so that the Euler-Lagrange equation (9.9) for the Schroedinger Lagrangian (9.16) becomes

$$i\hbar\,\frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V\,\psi, \tag{9.17}$$

where the Schroedinger equation for ψ^* is as an Euler-Lagrange equation (9.9) in terms of ψ, which yields

$$-i\hbar\,\frac{\partial \psi^*}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi^* + V\,\psi^*, \tag{9.18}$$

which is simply the complex-conjugate equation of Eq. (9.17).

The energy-momentum conservation law is now derived by Noether method. Because the potential $V(\mathbf{x}, t)$ is spatially nonuniform and time dependent, the energy-momentum contained in the wave field is not conserved and energy-momentum is exchanged between the wave field and the potential V. For example, the energy *transfer* equation is of the form

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla\cdot\mathbf{S} = |\psi|^2\,\frac{\partial V}{\partial t}, \tag{9.19}$$

where the energy density $\mathcal{E}$ and energy density flux $\mathbf{S}$ are given explicitly as

$$\begin{aligned}
\mathcal{E} &= -\mathcal{L}_R + \frac{i\hbar}{2}\left(\psi^*\,\frac{\partial \psi}{\partial t} - \psi\,\frac{\partial \psi^*}{\partial t}\right) \\
\mathbf{S} &= -\frac{\hbar^2}{2m}\left(\frac{\partial \psi}{\partial t}\,\nabla\psi^* + \frac{\partial \psi^*}{\partial t}\,\nabla\psi\right).
\end{aligned}$$

The momentum transfer equation, on the other hand, is

$$\frac{\partial \mathbf{P}}{\partial t} + \nabla\cdot\mathsf{T} = -|\psi|^2\,\nabla V, \tag{9.20}$$

where the momentum density $\mathbf{P}$ and momentum density tensor T are given explicitly as

$$\begin{aligned}
\mathbf{P} &= \frac{i\hbar}{2}\,(\psi\,\nabla\psi^* - \psi^*\,\nabla\psi) \\
\mathsf{T} &= \mathcal{L}_R\,\mathsf{I} + \frac{\hbar^2}{2m}\,(\nabla\psi^*\,\nabla\psi + \nabla\psi\,\nabla\psi^*).
\end{aligned}$$

Note that Eqs. (9.19) and (9.20) are exact equations.

Whereas energy-momentum is transfered between the wave field ψ and potential V, the amount of wave-action contained in the wave field is conserved. Indeed the wave-action conservation law is

$$\frac{\partial \mathcal{J}}{\partial t} \; + \; \nabla \cdot \mathbf{J} = 0, \tag{9.21}$$

where the wave-action density $\mathcal{J}$ and wave-action density flux $\mathbf{J}$ are

$$\mathcal{J} \; = \; \hbar \, |\psi|^2 \quad \text{and} \quad \mathbf{J} \; = \; \frac{i\hbar^2}{2m} \, (\psi \, \nabla \psi^* \; - \; \psi^* \, \nabla \psi) \tag{9.22}$$

Thus wave-action conservation law is none other than the law of conservation of probability associated with the normalization condition

$$\int \; |\psi|^2 \, d^3x \; = \; 1$$

for bounds states or the conservation of the number of quanta in a scattering problem.

9.4 Variational Principle for Maxwell's Equations*

The Lagrangian density for the evolution of electromagnetic fields in the presence of a charged-particle distribution is

$$\mathcal{L} \; = \; \frac{1}{8\pi} \left\{ \; |\mathbf{E}(\mathbf{x},t)|^2 \; - \; |\mathbf{B}(\mathbf{x},t)|^2 \; \right\} \; + \; \frac{1}{c} \, \mathbf{A}(\mathbf{x},t) \cdot \mathbf{J}(\mathbf{x},t) \; - \; \Phi(\mathbf{x},t) \, \rho(\mathbf{x},t), \tag{9.23}$$

where $\rho(\mathbf{x},t)$ and $\mathbf{J}(\mathbf{x},t)$ are the charge and current densities. The electric field $\mathbf{E}$ and magnetic field $\mathbf{B}$ are expressed in terms of the electromagnetic potentials Φ and $\mathbf{A}$ as

$$\mathbf{E} \; = \; - \, \nabla \Phi \; - \; \frac{1}{c} \, \frac{\partial \mathbf{A}}{\partial t} \quad \text{and} \quad \mathbf{B} \; = \; \nabla \times \mathbf{A}. \tag{9.24}$$

Note that as a result of the definitions (9.24), the electric and magnetic fields satisfy the conditions

$$\nabla \cdot \mathbf{B} \; = \; 0 \quad \text{and} \quad \nabla \times \mathbf{E} \; = \; - \, \frac{1}{c} \, \frac{\partial \mathbf{B}}{\partial t}, \tag{9.25}$$

which represent Gauss' Law for magnetic fields and Faraday's Law, respectively.

9.4.1 Maxwell's Equations as Euler-Lagrange Equations

The remaining Maxwell equations (Gauss's Law for electric fields – or Poisson's equation – and Ampère's Law) are derived as Euler-Lagrange equations for Φ and $\mathbf{A}$ as follows. The Euler-Lagrange equation for Φ is

$$\frac{\partial}{\partial t} \left\{ \; \frac{\partial \mathcal{L}}{\partial (\partial_t \Phi)} \; \right\} \; + \; \nabla \cdot \left\{ \; \frac{\partial \mathcal{L}}{\partial (\nabla \Phi)} \; \right\} \; - \; \frac{\partial \mathcal{L}}{\partial \Phi} \; = \; 0. \tag{9.26}$$

Here, $\partial\mathcal{L}/\partial(\partial_t\Phi) = 0$, $\partial\mathcal{L}/\partial\Phi = -\rho$, and

$$\frac{\partial\mathcal{L}}{\partial(\nabla\Phi)} \;=\; \frac{\partial}{\partial(\nabla\Phi)}\left\{\;\frac{1}{8\pi}\left(|\nabla\Phi + c^{-1}\,\partial_t\mathbf{A}|^2 - |\nabla\times\mathbf{A}|^2\right)\;\right\} \;=\; -\,\frac{\mathbf{E}}{4\pi}.$$

Hence, the Euler-Lagrange equation (9.26) becomes Gauss's Law for electric fields

$$\rho \;-\; \frac{\nabla\cdot\mathbf{E}}{4\pi} \;=\; 0. \tag{9.27}$$

The Euler-Lagrange equation for A_x is

$$\frac{\partial}{\partial t}\left\{\;\frac{\partial\mathcal{L}}{\partial(\partial_t A)_x)}\;\right\} \;+\; \nabla\cdot\left\{\;\frac{\partial\mathcal{L}}{\partial(\nabla A_x)}\;\right\} \;-\; \frac{\partial\mathcal{L}}{\partial A_x} \;=\; 0. \tag{9.28}$$

Here, $\partial\mathcal{L}/\partial A_x = J^x/c$,

$$\frac{\partial\mathcal{L}}{\partial(\partial_t A_x)} \;=\; \frac{\partial}{\partial(\partial_t A_x)}\left\{\;\frac{1}{8\pi}\left(|\nabla\Phi + c^{-1}\,\partial_t\mathbf{A}|^2 - |\nabla\times\mathbf{A}|^2\right)\;\right\} \;=\; -\,\frac{E^x}{4\pi\,c},$$

and

$$\begin{aligned}
\frac{\partial\mathcal{L}}{\partial(\partial_x A_x)} \;&=\; \frac{\partial}{\partial(\partial_x A_x)}\left\{\;\frac{1}{8\pi}.\left(|\nabla\Phi + c^{-1}\,\partial_t\mathbf{A}|^2 - |\nabla\times\mathbf{A}|^2\right)\;\right\} \;=\; 0,\\
\frac{\partial\mathcal{L}}{\partial(\partial_y A_x)} \;&=\; \frac{\partial}{\partial(\partial_y A_x)}\left\{\;\frac{1}{8\pi}\left(|\nabla\Phi + c^{-1}\,\partial_t\mathbf{A}|^2 - |\nabla\times\mathbf{A}|^2\right)\;\right\} \;=\; \frac{B^z}{4\pi},\\
\frac{\partial\mathcal{L}}{\partial(\partial_z A_x)} \;&=\; \frac{\partial}{\partial(\partial_z A_x)}\left\{\;\frac{1}{8\pi}\left(|\nabla\Phi + c^{-1}\,\partial_t\mathbf{A}|^2 - |\nabla\times\mathbf{A}|^2\right)\;\right\} \;=\; -\,\frac{B^y}{4\pi},
\end{aligned}$$

where

$$|\mathbf{B}|^2 \;=\; |\nabla\times\mathbf{A}|^2 \;=\; (\partial_y A_z - \partial_z A_y)^2 \;+\; (\partial_z A_x - \partial_x A_z)^2 \;+\; (\partial_x A_y - \partial_y A_x)^2.$$

Hence, the Euler-Lagrange equation (9.28) becomes

$$-\,\frac{1}{4\pi\,c}\,\frac{\partial E^x}{\partial t} \;+\; \frac{1}{4\pi}\left(\frac{\partial B^z}{\partial y} \;-\; \frac{\partial B^y}{\partial z}\right) \;-\; \frac{J^x}{c} \;=\; 0.$$

By combining the Euler-Lagrange equations for all three components of the vector potential $\mathbf{A}$, we thus obtain Maxwell's generalization of Ampère's Law

$$-\,\frac{1}{4\pi\,c}\,\frac{\partial\mathbf{E}}{\partial t} \;+\; \frac{\nabla\times\mathbf{B}}{4\pi} \;-\; \frac{\mathbf{J}}{c} \;=\; 0. \tag{9.29}$$

9.4.2 Energy Conservation Law for Electromagnetic Fields

We now derive The Noether equation for this Lagrangian density can be written as

$$\delta\mathcal{L} \;=\; -\,\frac{\partial}{\partial t}\left(\frac{\delta\mathbf{A}\cdot\mathbf{E}}{4\pi\,c}\right) \;-\; \nabla\cdot\left(\frac{\delta\Phi\;\mathbf{E}}{4\pi} \;+\; \frac{\delta\mathbf{A}\times\mathbf{B}}{4\pi}\right).$$

Use the Noether equation with $\delta\Phi = -\,\delta t\;\partial_t\Phi$, $\delta\mathbf{A} = -\,\delta t\;\partial_t\mathbf{A} = c\delta t\;(\mathbf{E}+\nabla\Phi)$, and

$$\delta\mathcal{L} \;=\; -\,\delta t\left\{\;\frac{\partial\mathcal{L}}{\partial t} \;-\; \left(\frac{\partial\mathcal{L}}{\partial t}\right)_{\Phi,\mathbf{A}}\right\},$$

where $(\partial_t\mathcal{L})_{\Phi,\mathbf{A}}$ denotes the explicit time dependence of the Lagrangian density (at constant Φ and $\mathbf{A}$), and the Euler-Lagrange equations derived in Part (a) to obtain the energy conservation law in the form

$$\frac{\partial\mathcal{E}}{\partial t} \;+\; \nabla\cdot\mathbf{S} \;=\; -\,\mathbf{E}\cdot\mathbf{J},$$

where

$$\mathcal{E} \;=\; \frac{1}{8\pi}\left(|\mathbf{E}|^2 \;+\; |\mathbf{B}|^2\right) \quad\text{and}\quad \mathbf{S} \;=\; \frac{c}{4\pi}\,\mathbf{E}\times\mathbf{B}$$

denote the energy density and the energy-density flux, respectively.

Using $\delta\Phi = -\,\delta t\;\partial_t\Phi$, $\delta\mathbf{A} = c\delta t\;(\mathbf{E}+\nabla\Phi)$, and

$$\delta\mathcal{L} \;=\; -\,\delta t\left\{\;\frac{\partial\mathcal{L}}{\partial t} \;-\; \left(\frac{\partial\mathcal{L}}{\partial t}\right)_{\Phi,\mathbf{A}}\right\},$$

the Noether equation becomes

$$-\,\frac{\partial\mathcal{L}}{\partial t} \;+\; \left(\frac{\partial\mathcal{L}}{\partial t}\right)_{\Phi,\mathbf{A}} \;=\; -\,\frac{\partial}{\partial t}\left\{(\mathbf{E}+\nabla\Phi)\cdot\frac{\mathbf{E}}{4\pi}\right\} \;-\; \nabla\cdot\left\{-\,\partial_t\Phi\;\frac{\mathbf{E}}{4\pi} \;+\; c\,(\mathbf{E}+\nabla\Phi)\times\frac{\mathbf{B}}{4\pi}\right\},$$

where

$$\left(\frac{\partial\mathcal{L}}{\partial t}\right)_{\Phi,\mathbf{A}} \;=\; \frac{\mathbf{A}}{c}\cdot\frac{\partial\mathbf{J}}{\partial t} \;-\; \Phi\;\frac{\partial\rho}{\partial t}.$$

By re-arranging terms, we find

$$\begin{aligned}&\frac{\partial}{\partial t}\left\{\frac{|\mathbf{E}|^2}{4\pi} \;-\; \frac{1}{8\pi}\left(|\mathbf{E}|^2-|\mathbf{B}|^2\right) \;-\; \frac{1}{c}\,\mathbf{A}\cdot\mathbf{J} \;+\; \Phi\,\rho \;+\; \nabla\Phi\cdot\frac{\mathbf{E}}{4\pi}\right\}\\ &+\;\nabla\cdot\left\{\frac{c}{4\pi}\,\mathbf{E}\times\mathbf{B} \;-\; \frac{\partial\Phi}{\partial t}\;\frac{\mathbf{E}}{4\pi} \;+\; \nabla\Phi\times\frac{c\,\mathbf{B}}{4\pi}\right\} \;=\; -\,\frac{\mathbf{A}}{c}\cdot\frac{\partial\mathbf{J}}{\partial t} \;+\; \Phi\;\frac{\partial\rho}{\partial t},\end{aligned}$$

or

$$\frac{\partial}{\partial t}\left[\frac{1}{8\pi}\left(|\mathbf{E}|^2+|\mathbf{B}|^2\right)\right] - \frac{1}{c}\left(\frac{\partial \mathbf{A}}{\partial t}\cdot\mathbf{J} + \mathbf{A}\cdot\frac{\partial \mathbf{J}}{\partial t}\right) + \left(\frac{\partial \Phi}{\partial t}\,\rho + \Phi\,\frac{\partial \rho}{\partial t}\right)$$
$$+ \frac{1}{4\pi}\left(\nabla\frac{\partial \Phi}{\partial t}\cdot\mathbf{E} + \nabla\Phi\cdot\frac{\partial \mathbf{E}}{\partial t}\right) + \nabla\cdot\left(\frac{c}{4\pi}\,\mathbf{E}\times\mathbf{B}\right) - \left(\nabla\frac{\partial \Phi}{\partial t}\cdot\frac{\mathbf{E}}{4\pi} + \frac{\partial \Phi}{\partial t}\,\frac{\nabla\cdot\mathbf{E}}{4\pi}\right)$$
$$+ \nabla\cdot\left[\nabla\times\left(c\,\Phi\frac{\mathbf{B}}{4\pi}\right) - \frac{c\,\Phi}{4\pi}\,\nabla\times\mathbf{B}\right] = -\frac{\mathbf{A}}{c}\cdot\frac{\partial \mathbf{J}}{\partial t} + \Phi\,\frac{\partial \rho}{\partial t}.$$

By defining the energy density and the energy-density flux

$$\mathcal{E} = \frac{1}{8\pi}\left(|\mathbf{E}|^2 + |\mathbf{B}|^2\right) \quad \text{and} \quad \mathbf{S} = \frac{c}{4\pi}\,\mathbf{E}\times\mathbf{B},$$

respectively, and performing some cancellations, we now obtain

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla\cdot\mathbf{S} = -\,\mathbf{E}\cdot\mathbf{J} - \frac{\nabla\Phi}{4\pi}\cdot\left(4\pi\,\mathbf{J} + \frac{\partial \mathbf{E}}{\partial t} - c\,\nabla\times\mathbf{B}\right) - \frac{\partial \Phi}{\partial t}\left(\rho - \frac{\nabla\cdot\mathbf{E}}{4\pi}\right).$$

Lastly, by substituting Eqs. (9.27) and (9.29), we obtain

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla\cdot\mathbf{S} = -\,\mathbf{E}\cdot\mathbf{J}.$$

The Noether method can also be used to derive an electromagnetic momentum conservation law and an electromagnetic wave action conservation law.

Appendix A

Notes on Feynman's Quantum Mechanics

A.1 Feynman postulates and quantum wave function

Feynman[1] makes use of the Principle of Least Action to derive the Schroedinger equation by, first, introducing the following postulates.

Postulate I: Consider the initial and final states a and b of a quantum system and the set $\mathcal{M}$ of all paths connecting the two states. The conditional probability amplitude $K(b|a)$ of finding the system in the final state b if the system was initially in state a is expressed as

$$K(b|a) \equiv \sum_{\mathcal{M}} \phi[X], \tag{A.1}$$

where the summation is over all paths $X(t)$ from a to b and the partial conditional probability amplitude associated with path $X(t)$ is

$$\phi[X] \propto \exp\left(\frac{i}{\hbar} S[X]\right), \tag{A.2}$$

with $S[X]$ corresponding to the *classical* action for this path. If we take the two states to be infinitesimally close to each other, i.e., whenever the points $a \equiv x_a$ (at time t_a) and $b \equiv x_a + \Delta x$ (at time $t_a + \Delta t$) are infinitesimally close, Eqs. (A.1)-(A.2) yield

$$K(x_a + \Delta x, t_a + \Delta t | x_a, t_a) = \frac{1}{A} \exp \frac{i}{\hbar} \left[\frac{m(\Delta x)^2}{2\,\Delta t} - \Delta t\, U\left(x_a + \frac{\Delta x}{2}, t_a + \frac{\Delta t}{2}\right) \right], \tag{A.3}$$

[1]The material presented in this Appendix is adapted from the book *Quantum Mechanics and Path Integrals* by R.P. Feynman and A.R. Hibbs (McGraw-Hill, New York, 1965).

where A is a normalization constant and the classical action integral

$$S[x] \;=\; \int_{t_a}^{t_a+\Delta t} \left[\frac{m}{2} \left(\frac{dx}{dt}\right)^2 \;-\; U(x,t) \right] dt \;=\; \frac{m(\Delta x)^2}{2\,\Delta t} \;-\; U\left(x_a + \frac{\Delta x}{2}, t_a + \frac{\Delta t}{2}\right)\, \Delta t$$

is a function of Δx and Δt as well as the initial space-time point (x_a, t_a); here, the potential energy U is evaluated at the mid-point between x_a and x_b at a time $t_a + \Delta t/2$.

Postulate I provides the appropriate explanation for the mystery behind the Principle of Least Action ("*act locally, think globally*"). Indeed, in the classical limit ($\hbar \to 0$), we find that the variations δX around the paths $X(t)$ which are far away from the physical path $x(t)$ yield changes δS for which $\delta S/\hbar$ is large. Consequently, such contributions tend to average out to zero because of the corresponding wild oscillations in $\phi[X]$. On the other hand, for variations δX near the physical path $x(t)$, we get $\delta S \sim 0$ (to first order) and, consequently, paths $X(t)$ for which $S[X]$ is within $\hbar$ of S_{cl} will contribute strongly. The resulting effect is that only paths in the neighborhood of the physical path $x(t)$ have a nonvanishing probability amplitude. In the strict limit $\hbar \to 0$, the only such path with a nonvanishing probability amplitude is the physical path.

Postulate II: The quantum wave function $\psi(x,t)$ is defined as the probability amplitude for the particle to be at the location x at time t, i.e., $\psi(x,t) \equiv K(x,t|\bullet)$, where we are not interested in the previous history of the particle (its previous location is denoted by $\bullet$) but only on its future time evolution. The Second Postulate states that the integral equation relating the wave function $\psi(x_2,t_2)$ to the wave function $\psi(x_1,t_1)$ is given as

$$\psi(x_2,t_2) \;\equiv\; \int_{-\infty}^{\infty} dx_1\; K(x_2,t_2|x_1,t_1)\; \psi(x_1,t_1). \tag{A.4}$$

If we set in Eq. (A.4): $t_1 = t$ and $t_2 = t + \epsilon$, $x_2 = x$ and $x_1 = x + \eta$, then for small enough values of ϵ, the conditional probability amplitude (A.3) can be used in (A.4) to yield

$$\begin{aligned} \psi(x,t+\epsilon) \;&\equiv\; \int_{-\infty}^{\infty} d\eta\; K(x,t+\epsilon|x+\eta,t)\; \psi(x+\eta,t) \qquad (A.5)\\ &=\; \int_{-\infty}^{\infty} \frac{d\eta}{A}\; \exp\left[\, \frac{i}{\hbar} \left(\frac{m\eta^2}{2\epsilon} \;-\; \epsilon\, U(x+\eta/2,\, t+\epsilon/2) \right) \right] \psi(x+\eta,t), \end{aligned}$$

where a time-dependent potential $U(x,t)$ is considered.

A.2 Derivation of the Schroedinger equation

We will now derive the Schroedinger equation by expanding both sides of Eq. (A.5), up to first order in ϵ (and neglect all higher powers). First, on the left side of Eq. (A.5), we have

$$\psi(x,t+\epsilon) \;=\; \psi(x,t) \;+\; \epsilon\, \frac{\partial \psi(x,t)}{\partial t}. \tag{A.6}$$

Next, on the right side of Eq. (A.5), we note that the exponential

$$\exp\left[\left(\frac{im}{2\hbar}\right)\frac{\eta^2}{\epsilon}\right]$$

oscillates wildly as $\epsilon \to 0$ for all values of η except those for which $m\eta^2/(2\hbar\epsilon) \sim 1$. We, therefore, conclude that the contribution from the integral in Eq. (A.5) will come from values $\eta = \mathcal{O}(\epsilon^{1/2})$ and, consequently, we may expand the remaining functions appearing on the right side of Eq. (A.5) up to η^2. Thus, we may write

$$\psi(x+\eta,t) \;=\; \psi(x,t) \;+\; \eta\,\frac{\partial\psi(x,t)}{\partial x} \;+\; \frac{\eta^2}{2}\,\frac{\partial^2\psi(x,t)}{\partial x^2},$$

while

$$\exp\left[-\frac{i\epsilon}{\hbar}\,U(x+\eta/2,\,t+\epsilon/2)\right] \;=\; 1 \;-\; \frac{i\epsilon}{\hbar}\,U(x,t).$$

Expanding the right side of Eq. (A.5) to first order in ϵ, therefore, yields

$$\left[1 \;-\; \frac{i\epsilon}{\hbar}\,U(x,t)\right]\psi(x,t) \;+\; \left(\frac{I_1}{I_0}\right)\frac{\partial\psi(x,t)}{\partial x} \;+\; \left(\frac{I_2}{2\,I_0}\right)\,\frac{\partial^2\psi(x,t)}{\partial x^2}, \tag{A.7}$$

where $A = I_0$ and

$$I_n \;\equiv\; \int_{-\infty}^{\infty} d\eta\;\eta^n\;e^{-a\eta^2}, \quad a \equiv \frac{m}{2i\hbar\epsilon}$$

with $I_0 = \sqrt{\pi/a}$, $I_1 = 0$, and $I_2 = 1/2a = (i\hbar/m)\,\epsilon$.

Lastly, we find that the terms of first order in ϵ in Eqs. (A.6) and (A.7) must be equal, and we obtain

$$\frac{\partial\psi(x,t)}{\partial t} \;=\; -\,\frac{i}{\hbar}\,U(x,t)\;\psi(x,t) \;+\; \frac{i\hbar}{2m}\,\frac{\partial^2\psi(x,t)}{\partial x^2},$$

or

$$i\hbar\,\frac{\partial\psi(x,t)}{\partial t} \;=\; -\,\frac{\hbar^2}{2m}\,\frac{\partial^2\psi(x,t)}{\partial x^2} \;+\; U(x,t)\;\psi(x,t). \tag{A.8}$$

This equation is known as the Schroedinger equation and it describes the time evolution of the wave function $\psi(x,t)$.