Classical Mechanics

Jung Hoon Han

May 18, 2015

Contents

Chapter 1

Coordinate Systems

1.1 Orthogonal Rotation in Cartesian Coordinates

Analytic description of classical motion rests on the definition of coordinate system within which the motion is described. The most frequently used one is the Cartesian system, labeled by the three real numbers indicating the component of a vector along the x, y, and z axes. Mathematically, we write a vector in terms of its components as

$$\mathbf{v} = v_x\hat{x} + v_y\hat{y} + v_z\hat{z}. \tag{1.1}$$

The meaning of the above notation is that a vector $\mathbf{v}$ is characterized by three numbers (v_x, v_y, v_z) each expressing the amount of steps taken from the origin along a particular axis to form that vector. The three axes, $\{\hat{x}, \hat{y}, \hat{z}\}$ are assumed to be orthogonal,

$$\hat{x} \cdot \hat{y} = \hat{y} \cdot \hat{z} = \hat{z} \cdot \hat{x} = 0. \tag{1.2}$$

The direction vectors themselves are of course vectors, whose lengths are taken to be unity: $\hat{x} \cdot \hat{x} = \hat{y} \cdot \hat{y} = \hat{z} \cdot \hat{z} = 1$. In this way, we arrive at an orthonormal system of basis vectors, $\hat{e}_\alpha$, where $\alpha = 1, 2, 3$ run over the three unit vectors with the property

$$\hat{e}_\alpha \cdot \hat{e}_\beta = \delta_{\alpha\beta}. \tag{1.3}$$

We use the Kronecker delta symbol $\delta_{\alpha\beta}$ for a function $+1$ when $\alpha = \beta$, and 0 when $\alpha \neq \beta$.

One can easily imagine the coordinate frame rotated as a rigid body, to arrive at a new set of basis vectors $\{\hat{e}'_\alpha\}$. In the new frame, the same vector $\mathbf{v}$ must be assigned different set of coordinates v'_α,

$$\mathbf{v} = \sum_\alpha v_\alpha \hat{e}_\alpha = \sum_\alpha v'_\alpha \hat{e}'_\alpha. \tag{1.4}$$

Furthermore, the two set of components $\{v_\alpha\}$ and $\{v'_\alpha\}$ must be related to each other.

To see their relationship, take the inner product of both sides with $\hat{e}_\beta$, one of the unit vectors in the original coordinate frame. By utilizing the Kronecker relation, we see that the l.h.s. only gives v_β, which the r.h.s. gives a set of numbers added together $\sum_\alpha v'_\alpha(\hat{e}'_\alpha \cdot \hat{e}_\beta)$. The inner product of unit vectors in the new and the old coordinate frames are called direction cosines: $\hat{e}'_\alpha \cdot \hat{e}_\beta = \lambda_{\beta\alpha}$. Now the relation is neatly summarized as

$$v_\beta = \sum_\alpha \lambda_{\beta\alpha} v'_\alpha. \tag{1.5}$$

If we arrange the components v_β as a column vector $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ and $\lambda_{\beta\alpha}$ as a matrix, $\boldsymbol{\lambda} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{pmatrix}$, Eq. (1.5) is nothing but a matrix equation

$$\mathbf{v} = \boldsymbol{\lambda}\mathbf{v}'. \tag{1.6}$$

Components of the same vector in the new and old coordinate frames are related by a linear matrix relationship like the one shown above. If we take the inner product of Eq. (1.4) with respect to $\hat{e}'_\beta$, we would arrive instead at

$$\sum_\alpha v_\alpha(\hat{e}_\alpha \cdot \hat{e}'_\beta) = v'_\beta. \tag{1.7}$$

Invoking the definition of direction cosines, it becomes $\sum_\alpha v_\alpha \lambda_{\alpha\beta} = v'_\beta$, or

$$\mathbf{v}' = \boldsymbol{\lambda}^T \mathbf{v}. \tag{1.8}$$

The uppercase T denotes the transpose of a matrix. On the other hand, we can obtain the same relationship by multiplying Eq. (1.6) by the inverse matrix of $\boldsymbol{\lambda}$:

$$\mathbf{v}' = \boldsymbol{\lambda}^{-1}\mathbf{v}. \tag{1.9}$$

By comparing the two equations, (1.8) and (1.9) we are forced to conclude that $\boldsymbol{\lambda}$ obeys a special property

$$\boldsymbol{\lambda}^T = \boldsymbol{\lambda}^{-1}. \tag{1.10}$$

The transpose of the direction cosine matrix is its inverse. Such matrices are known as orthogonal matrices. The condition $\boldsymbol{\lambda}^T\boldsymbol{\lambda}^T = 1$ implies

$$\lambda_{ij}\lambda_{ik} = \delta_{jk}. \tag{1.11}$$

It looks somewhat odd, but in fact there is a well-defined geometric meaning to this identity. Note that $\lambda_{ij} = \hat{e}_i \cdot \hat{e}'_j$. Substituting, we realize that the identity above means

$$(\hat{e}_i \cdot \hat{e}'_j)(\hat{e}_i \cdot \hat{e}'_k) = \delta_{jk}. \tag{1.12}$$

Where does this relation come from? Since $\hat{e}'_j = (\hat{e}'_j \cdot \hat{e}_i)\hat{e}_i$, and $\hat{e}'_k = (\hat{e}'_k \cdot \hat{e}_i)\hat{e}_i$, their inner product, $\hat{e}'_j \cdot \hat{e}'_k$, becomes $(\hat{e}'_j \cdot \hat{e}_i)(\hat{e}'_k \cdot \hat{e}_i)$. On the other hand, $\hat{e}'_j \cdot \hat{e}'_k = \delta_{jk}$, so we obtain Eq. (1.11).

Once we know the direction cosine matrix $\boldsymbol{\lambda}$, we can map the coordinates from one frame to any other by the linear multiplication of the matrix.

In two dimensions the direction cosine matrix is 2×2 given by

$$\boldsymbol{\lambda} = \begin{pmatrix} \hat{e}_1 \cdot \hat{e}'_1 & \hat{e}_1 \cdot \hat{e}'_2 \\ \hat{e}_2 \cdot \hat{e}'_1 & \hat{e}_2 \cdot \hat{e}'_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \tag{1.13}$$

We arrive at the familiar results

$$\begin{aligned} x &= x' \cos\theta - y' \sin\theta \\ y &= y' \cos\theta + x' \sin\theta. \end{aligned} \tag{1.14}$$

Determinant of a transpose $\boldsymbol{\lambda}^T$ is the same as the determinant itself, $\mathrm{Det}(\boldsymbol{\lambda}) = \mathrm{Det}(\boldsymbol{\lambda}^T)$. On the other hand, determinant of an inverse matrix is the inverse of the determinant: $\mathrm{Det}(\boldsymbol{\lambda}^{-1}) = 1/\mathrm{Det}(\boldsymbol{\lambda})$. Since $\boldsymbol{\lambda}^T = \boldsymbol{\lambda}^{-1}$, we must conclude $\mathrm{Det}(\boldsymbol{\lambda}) = 1/\mathrm{Det}(\boldsymbol{\lambda})$, or $\mathrm{Det}(\boldsymbol{\lambda}) = \pm 1$. Those frame changes with $\mathrm{Det}(\boldsymbol{\lambda}) = +1$ are called "proper" rotations, while those with $\mathrm{Det}(\boldsymbol{\lambda}) = -1$ are "improper". A notable example of improper frame change is an inversion, which takes one of the unit vectors to its minus. For instance, $\hat{x} \to -\hat{x}$, $\hat{y} \to \hat{y}$, $\hat{z} \to \hat{z}$ is an inversion. The direction cosine matrix is $\boldsymbol{\lambda} = \mathrm{Diag}(-1, 1, 1)$ and its determinant is -1. Similarly, reversing all three axes ($\hat{x} \to -\hat{x}$, $\hat{y} \to -\hat{y}$, $\hat{z} \to -\hat{z}$) is improper, but reversing only two of the axes is proper ($\hat{x} \to -\hat{x}$, $\hat{y} \to -\hat{y}$, $\hat{z} \to \hat{z}$). Indeed, the last example is simply the rotation about the $\hat{z}$-axis by the angle π.

Successive changes of the coordinate frames are represented as a product of two orthogonal matrices $\boldsymbol{\lambda}_1\boldsymbol{\lambda}_2$. If we implement the rotations in the opposite order, we generally get a different result: $\boldsymbol{\lambda}_2\boldsymbol{\lambda}_1 \neq \boldsymbol{\lambda}_1\boldsymbol{\lambda}_2$. Objects which do not coommute are sometimes called non-Abelian. Matrices are examples of non-Abelian objects. If we restrict ourselves to rotations about the same axis, say the $\hat{z}$-axis, then the matrices begin to commute again. Write $\lambda(\theta)$ for a rotation about the $\hat{z}$-axis by angle θ, then one can easily show by explicit matrix product calculation that

$$\boldsymbol{\lambda}(\theta_1)\boldsymbol{\lambda}(\theta_2) = \boldsymbol{\lambda}(\theta_1 + \theta_2) = \boldsymbol{\lambda}(\theta_2)\boldsymbol{\lambda}(\theta_1). \tag{1.15}$$

The subset of rotation matrices about the common axis is Abelian.

Two vectors $\mathbf{A}$ and $\mathbf{B}$ can be combined to produce a scalar (by taking inner product $c = \mathbf{A} \cdot \mathbf{B}$) or another vector (by taking an outer product $\mathbf{C} = \mathbf{A} \times \mathbf{B}$). Let's examine how the two quantities transform under the orthogonal rotation.

In the new frame, we have $c' = \mathbf{A}' \cdot \mathbf{B}'$. According to the transformation rule $\mathbf{A}' = \boldsymbol{\lambda}\mathbf{A}$ and $\mathbf{B}' = \boldsymbol{\lambda}\mathbf{B}$, the new inner product becomes

$$c' = (\boldsymbol{\lambda}\mathbf{A}) \cdot (\boldsymbol{\lambda}\mathbf{B}) = (\mathbf{A} \cdot \mathbf{B}) = c. \tag{1.16}$$

For explicit proof,

$$c' = A'_i B'_i = (R_{ij}A_j)(R_{ik}B_k) = (R_{ij}R_{ik})A_j B_k = \delta_{jk}A_j B_k = A_j B_j = \mathbf{A} \cdot \mathbf{B}. \tag{1.17}$$

The inner product remains unchanged under the coordinate rotation. Such quantities are called scalars. On the other hand, $\mathbf{C}' = \mathbf{A}' \times \mathbf{B}'$ is given by

$$\mathbf{C}' = (\boldsymbol{\lambda}\mathbf{A}) \times (\boldsymbol{\lambda}\mathbf{B}) = \boldsymbol{\lambda}(\mathbf{A} \times \mathbf{B}) = \boldsymbol{\lambda}\mathbf{C}. \tag{1.18}$$

Given the transformation property, $\mathbf{C}$ is a vector, which of course it is. To prove it, first work out both $\boldsymbol{\lambda}(\mathbf{A} \times \mathbf{B})$ and $(\boldsymbol{\lambda}\mathbf{A}) \times (\boldsymbol{\lambda}\mathbf{B})$:

$$\begin{aligned} [\boldsymbol{\lambda}(\mathbf{A} \times \mathbf{B})]_i &= \lambda_{ij}(\mathbf{A} \times \mathbf{B})_j = \lambda_{ij}\varepsilon_{jmn}A_m B_n, \\ [(\boldsymbol{\lambda}\mathbf{A}) \times (\boldsymbol{\lambda}\mathbf{B})]_i &= \varepsilon_{ijk}(\boldsymbol{\lambda}\mathbf{A})_j(\boldsymbol{\lambda}\mathbf{B})_k = \varepsilon_{ijk}\lambda_{jm}\lambda_{kn}A_m B_n. \end{aligned} \tag{1.19}$$

To prove the equality we need to show $\lambda_{ij}\varepsilon_{jmn} = \varepsilon_{ijk}\lambda_{jm}\lambda_{kn}$. Multiplying both sides by λ_{ip} and summing over i gives

$$\varepsilon_{pmn} = \varepsilon_{ijk}\lambda_{ip}\lambda_{jm}\lambda_{kn}. \tag{1.20}$$

The r.h.s. is the determinant of a matrix consisting of three columns p, m, n of the λ matrix. If any two columns coincide, the determinant is zero. For all three columns non-overlapping, we get either +1 or -1 depending on the order the column vectors are arranged. Proof is complete.

Two useful vector identities are listed below.

$$\begin{aligned} \mathbf{A} \times (\mathbf{B} \times \mathbf{C}) &= \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B}), \\ (\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) &= (\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D})(\mathbf{B} \cdot \mathbf{C}). \end{aligned} \tag{1.21}$$

The first one can be proven with the identity $\varepsilon_{ijk}\varepsilon_{ilm} = \delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}$. Using similar ideas, one can also show

$$\begin{aligned} \boldsymbol{\nabla} \times (\mathbf{B} \times \mathbf{C}) &= \mathbf{B}(\boldsymbol{\nabla} \cdot \mathbf{C}) - \mathbf{C}(\boldsymbol{\nabla} \cdot \mathbf{B}) + (\mathbf{C} \cdot \nabla)\mathbf{B} - (\mathbf{B} \cdot \nabla)\mathbf{C}, \\ \boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{C}) &= \boldsymbol{\nabla}(\boldsymbol{\nabla} \cdot \mathbf{C}) - \boldsymbol{\nabla}^2\mathbf{C}. \end{aligned} \tag{1.22}$$

1.2 Curved Coordinates

Sometimes we choose to describe the motion of a particle in cylindrical, or spherical coordinates. Instead of three orthonormal basis vectors which are fixed, the new coordinate frames employ basis vectors that are co-moving with the particle's trajectory itself. In cylindrical coordinates, the particle position $\mathbf{r}$ is expressed as

$$\mathbf{r} = r\hat{r} + z\hat{z}, \tag{1.23}$$

where the radial unit vector $\hat{r}$ follows from the definition

$$\hat{r} = (\cos\theta, \sin\theta). \tag{1.24}$$

The relation $(x, y) = r(\cos\theta, \sin\theta)$, where $r = \sqrt{x^2 + y^2}$ is the radial distance of the particle measured from the origin, easily establishes the identity of the cylindrical coordinate representation of the position vector $\mathbf{r}$ with that of the Cartesian system. Since the motion along the z-axis is trivial, we will confine ourselves to the two-dimensional motion in which $\mathbf{r} = r\hat{r}$.

The translation of the particle's coordinate from Cartesian to cylindrical system is purely kinetic. In other words they are just two mathematically equivalent ways of writing the same position vector $\mathbf{r}$. When we come to discuss Newtonian dynamics it is important to know how to express both the velocity and the acceleration vectors, given as first and second derivatives of the coordinate $\mathbf{r}$ respectively, in cylindrical coordinates.

Velocity and acceleration vectors in the Cartesian coordinate system are trivial to obtain as

$$\begin{aligned} \dot{\mathbf{r}} &= \dot{x}\hat{x} + \dot{y}\hat{y} + \dot{z}\hat{z}, \\ \ddot{\mathbf{r}} &= \ddot{x}\hat{x} + \ddot{y}\hat{y} + \ddot{z}\hat{z}, \end{aligned} \tag{1.25}$$

due to the fact that bases vectors $\hat{x}, \hat{y}, \hat{z}$ are independent of time. It is no longer the case that the basis vectors would remain time-dependent in other coordinate systems.

When we take the time derivative of $\mathbf{r}$ to obtain the velocity, not only do we need to worry about $\dot{r}$ but also the time derivative of $\hat{r}$. So let's calculate it.

$$\frac{d}{dt}\hat{r} = \frac{d}{dt}\frac{\mathbf{r}}{r} = \frac{\dot{\mathbf{r}}}{r} - \frac{\dot{r}\mathbf{r}}{r^2}. \tag{1.26}$$

This looks awkward. In fact there is a better way to express the time derivative of $\hat{r}$, if we first recall the identity $d(\hat{r}\cdot\hat{r})/dt = d(1)/dt = 0 = 2\hat{r}\cdot\dot{\hat{r}}$. So whatever $\dot{\hat{r}}$ is, it ought to be orthogonal to the original vector $\hat{r}$. In two dimensions, there is only one such direction. We call it

$$\hat{\theta} = (-\sin\theta, \cos\theta). \tag{1.27}$$

But still it's not very easy to prove, based on the above expression for $d\hat{r}/dt$, that it is orthogonal to $\hat{r}$ and parallel to $\hat{\theta}$. So we choose a different strategy and go back to the coordinate expressions of $\hat{r}$ and $\hat{\theta}$, which were

$$\hat{r} = (\cos\theta, \sin\theta), \quad \hat{\theta} = (-\sin\theta, \cos\theta). \tag{1.28}$$

Now it is quite obvious by direct differentiation that

$$\frac{d\hat{r}}{dt} = \dot{\theta}\hat{\theta}, \ \frac{d\hat{\theta}}{dt} = -\dot{\theta}\hat{r}. \tag{1.29}$$

Now the velocity is calculated easily:

$$\mathbf{v} = \dot{r}\hat{r} + r\frac{d\hat{r}}{dt} = \dot{r}\hat{r} + r\dot{\theta}\hat{\theta}. \tag{1.30}$$

Repeating the same steps, one can derive the acceleration vector in the cylindrical coordinates as well:

$$\mathbf{a} = (\ddot{r} - r\dot{\theta}^2)\hat{r} + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\hat{\theta}. \tag{1.31}$$

In the case of central forces, the acceleration has only the radial component while the tangential component must vanish. The vanishing quantity can be written as

$$r^2\ddot{\theta} + 2r\dot{r}\dot{\theta} = \frac{d}{dt}(r^2\dot{\theta}) = 0. \tag{1.32}$$

This is just a re-statement of Kepler's second law of planetary motion, or the conservation of angular momentum.

Another popular coordinate system used in the physical description is the spherical coordinates. Here all three basis vectors given below are co-moving with the particle's location.

$$\begin{aligned} \hat{r} &= (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \\ \hat{\theta} &= (\cos\theta\cos\phi, \cos\theta\sin\phi, -\sin\theta) \\ \hat{\phi} &= (-\sin\phi, \cos\phi, 0). \end{aligned} \tag{1.33}$$

The second unit vector $\hat{\theta}$ is obtained from the first by differentiation, $\partial_\theta \hat{r} = \hat{\theta}$. The third one is obtained as the cross product of the first two: $\hat{\phi} = \hat{r} \times \hat{\theta}$. Time derivative of each unit vector can be expressed as the linear combination of the other two since it has no component along its own basis vector.

Starting from $\mathbf{r} = r\hat{r}$, the velocity vector and the acceleration vector can be derived as

$$\begin{aligned} \mathbf{v} &= \dot{r}\hat{r} + r\dot{\theta}\hat{\theta} + r\sin\theta\dot{\phi}\hat{\phi}, \\ \mathbf{a} &= (\ddot{r} - r\dot{\phi}^2\sin^2\theta - r\dot{\theta}^2)\hat{r} \\ &+ (r\ddot{\theta} + 2\dot{r}\dot{\theta} - r\dot{\phi}^2\sin\theta\cos\theta)\hat{\theta} \\ &+ (r\ddot{\phi}\sin\theta + 2\dot{r}\dot{\phi}\sin\theta + 2r\dot{\theta}\dot{\phi}\cos\theta)\hat{\phi}. \end{aligned} \tag{1.34}$$

1.3 Problems

1. Prove Eq. (1.15).
2. Prove Eq. (1.16).
3. Prove Eq. (1.18).
4. Prove Eq. (1.21).
5. Prove Eq. (1.22).
6. Prove Eq. (1.31).
7. Prove Eq. (1.34). You will first have to show how to express $\dot{\hat{r}}$, $\dot{\hat{\theta}}$, and $\dot{\hat{\phi}}$ in terms of $\hat{r}, \hat{\theta}, \hat{\phi}$.
8. An earth revolves around the sun in a circle of radius R_s with angular velocity ω_s. On its own, the earth of radius R_e rotates about its own axis with angular speed ω_e. Find the acceleration at arbitrary point on the surface of the earth, using the sun as the origin of the coordinate system. For simplicity take the axis of earth's rotation to be perpendicular to the plane of earth's motion about the sun, instead of being tilted as it is in reality.

Chapter 2

Lagrangian and Hamiltonian

2.1 Lagrangian

Newton's force law states

$$m\frac{d\mathbf{v}}{dt} = \mathbf{F} = -\boldsymbol{\nabla} V, \tag{2.1}$$

assuming a conservative force $\mathbf{F} = -\boldsymbol{\nabla} V$ that can be derived from the scalar function (called the potential energy) through the operation of taking the gradient on it. Multiplying both sides of Newton's equation by $\mathbf{v}$ gives

$$\frac{d}{dt}\left(\frac{1}{2}m\mathbf{v}^2\right) = -\frac{d\mathbf{r}}{dt}\cdot\boldsymbol{\nabla} V = -\frac{dV(\mathbf{r})}{dt}, \tag{2.2}$$

which in turn implies the conservation of total energy

$$\frac{d}{dt}(K+V) = 0,\ K = \frac{1}{2}m\mathbf{v}^2. \tag{2.3}$$

Starting from Newton's force law, which is expressed in terms of vector quantities, the principle of energy conservation is derived following the steps outlined above. One might say that in this approach force is fundamental, while energy is the concept derived from the underlying notion of force.

Another school of thought takes the energy as the more fundamental quantity, from which Newton's law will be derived according to some firm principle of nature. Champions of such thoughts are Leibniz, Bernoulli, Lagrange, and Hamilton, among others.

To get a grasp of the Lagrangian approach to mechanics, it is essential to understand first the notion of a functional - a function of a function. Kinetic energy, $K = (m/2)\dot{\mathbf{r}}^2$, for instance, is a function of the velocity $\mathbf{v}$. But since the velocity itself is a function of the trajectory $\mathbf{r}(t)$ which depends on time t, we may say more correctly that the kinetic energy is a function of the "whole

history of the trajectory $\mathbf{r}(t)$" as a function of time t. The same argument can be made with the potential energy $V(\mathbf{r}(t))$. Let's write down the following quantity,

$$\mathcal{K} = \int_{t_1}^{t_2} K(\dot{\mathbf{r}}(t))dt, \qquad \mathcal{V} = \int_{t_1}^{t_2} V(\mathbf{r}(t))dt, \tag{2.4}$$

which as we argued is a functional of the trajectory, $\mathbf{r}(t)$. The difference of these two, $\mathcal{S} = \mathcal{K} - \mathcal{V}$, called the action, is also a functional of the trajectory $\mathbf{r}(t)$. The integrand $\mathcal{S} = K - V$ is called the Lagrangian. The action is then the time integral of the Lagrangian for a particular chosen path $\mathbf{r}(t)$.

Think of a particular trajectory $\mathbf{r}(t)$ as a piece of rubber stretched in some way in space. Next consider a trajectory which is just a slight displacement of the original trajectory, $\mathbf{r}'(t) = \mathbf{r}(t) + \delta\mathbf{r}(t)$. At each time t, $\delta\mathbf{r}(t)$ is assumed an infinitesimal quantity. What is the difference in the value of the action $\mathcal{S}$ adopted by the two trajectories? This is basically an enlarged version of the question: what is the difference in the value of the function $f(x)$ when x is displaced to $x + \delta x$? Such question led to the notion of differentiation as we know so well. A similar question can be posed for the variation in path, and the difference it causes in the functional $\mathcal{S}$. To answer the question, simply consider the values of $\mathcal{K}$ for the two paths and subtract:

$$\delta\mathcal{K} = \int_{t_1}^{t_2} [K(\mathbf{r} + \delta\mathbf{r}) - K(\mathbf{r})]dt \approx \int_{t_1}^{t_2} m\mathbf{v} \cdot \frac{d}{dt}\delta\mathbf{r}. \tag{2.5}$$

In going to the last expression we ignored terms of order $(\delta\mathbf{r})^2$. Integrating by parts, with the boundary conditions $\delta\mathbf{r}(t_1) = 0 = \delta\mathbf{r}(t_2)$, we arrive at

$$\delta\mathcal{K} = -\int_{t_1}^{t_2} \delta\mathbf{r} \cdot \frac{d}{dt}(m\mathbf{v}). \tag{2.6}$$

A similar exercise gives

$$\delta\mathcal{V} = \int_{t_1}^{t_2} \delta\mathbf{r} \cdot \boldsymbol{\nabla} V. \tag{2.7}$$

Variation in the action is

$$\delta\mathcal{S} = \delta\mathcal{K} + \delta\mathcal{V} = -\int_{t_1}^{t_2} dt\delta\mathbf{r} \cdot \left(\frac{d}{dt}(m\mathbf{v}) + \boldsymbol{\nabla} V\right). \tag{2.8}$$

The quantity inside parenthesis is exactly Newton's equation. If the original path we have chosen to vary by a tiny amount happened to coincide with the actual, physical trajectory obeying Newton's law, the corresponding action $\mathcal{S}$ will have a vanishing difference against the small variation in the path. In other words, the action will achieve the extremal value for the path actually realized in nature.

The spirit of the statement can be reversed: If a path is found such that it gives the extremal value of the action, then such path would be the physical path obeying Newton's law. Hence, the entire classical mechanics can be formulated in terms of the new quantity, the action $\mathcal{S}$, together with the variational principle. This we might state as an equally compelling formulation of the laws of classical mechanics.

We know that some problems in mechanics can be better tackled in curved coordinate frames. Sometimes it is better to describe the particle's trajectories in terms of angles and distances from the origin. It is fitting that we enlarge the notion of coordinates to refer to any set of variables being used to completely specify the particle's location at any given time. We call them q_i, $i = 1, \cdots, N$. The quantity $L = T - V$ is also now a function of the generalized coordinates q_i and their derivatives $\dot{q}_i$: $L(q_i, \dot{q}_i)$. Applying the variational principle for the action written in terms of the generalized coordinates and their derivatives results in

$$\delta\mathcal{S} = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q_i}\delta q_i + \frac{\partial L}{\partial \dot{q}_i}\delta\dot{q}_i \right) dt = \int_{t_1}^{t_2} \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i dt. \tag{2.9}$$

The final expression is once again obtained by doing integration by parts subject to the boundary condition $\delta q_i(t_1) = \delta q_i(t_2) = 0$. Assuming the variational principle really works, we obtain a set of equations governing the classical motion as

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = 0. \tag{2.10}$$

The Lagrange's equation can be used in place of Newton's force equation as the law of classical mechanics.

As an example of the Lagrangian and the variational principle at work, let us begin with the simplest example of one-dimensional simple harmonic oscillator (SHO), $K = (m/2)\dot{x}^2$, $V = (k/2)x^2$. You can derive Newton's equation $m\ddot{x} = -kx$ by applying Lagrange's equation in no time.

As a second, more challenging example, consider a simple pendulum of length l with a mass m attached at its end. The trajectory of a swinging mass spans the two dimensional plane, but apparently one only needs one coordinate θ (angle from the vertical position) to describe the motion. Somehow the reduction of the number of coordinates from two to one has to be accounted for. First write down the Lagrangian of a two-dimensional motion taking place in the xz plane, subject to a constant gravity $\mathbf{F} = -g\hat{z}$.

$$L = \frac{m}{2}(\dot{x}^2 + \dot{z}^2) + mgz. \tag{2.11}$$

This is not quite the whole story, since the two coordinates are constrained by the condition $x^2 + z^2 = l^2$ which implies, among other things, $x\dot{x} + z\dot{z} = 0$. One way to deal with the constraint is to introduce a new coordinate system in which the constraint is naturally removed. Take $x = l\sin\theta$, $z = l\cos\theta$, and $\dot{x} = \dot{\theta}\cos\theta$, $\dot{z} = -\dot{\theta}\sin\theta$. The Lagrangian becomes

$$L = \frac{ml^2}{2}\dot{\theta}^2 + mgl\cos\theta. \tag{2.12}$$

The variational principle then immediately gives out the equation of the pendulum motion:

$$ml^2\ddot{\theta} = -mgl\sin\theta \Rightarrow \ddot{\theta} = -\frac{g}{l}\sin\theta. \tag{2.13}$$

Restriction to the small-amplitude oscillation ($|\theta| \ll 1$) will recover the SHO: $\ddot{\theta} = -\omega^2\theta$, $\omega^2 = g/l$.

Another way to solve the problem of constrained motion such as that of a pendulum is to use a trick that bear Lagrange's name, called the Lagrange multipler method. For the pendulum problem the constraint was simple enough that we could solve it right away within the suitable new coordinates that did away with the constraint. The method will not work, however, if the constraint got too complicated. In this case the trick is to supplement the original action $\mathcal{S}$ by a term proportional to the constraint.

$$\mathcal{S} = \int_{t_1}^{t_2} dt\ \left(\frac{m}{2}(\dot{x}^2 + \dot{z}^2) + mgz\right) + \int_{t_1}^{t_2} dt\ \lambda(x^2 + z^2 - l^2). \tag{2.14}$$

The new variable λ, also time-dependent, is to be regarded as an additional coordinate, to be varied in the same manner as x and z. Variation of λ results in the Lagrange's equation

$$\frac{\partial L}{\partial \lambda} = x^2 + y^2 - \lambda^2 = 0, \tag{2.15}$$

which is nothing but the constraint. Now the constraint is implemented as part of an equation of motion. Other components of Lagrange's equations are

$$m\ddot{x} = 2\lambda x,\ m\ddot{y} = 2\lambda y + mg. \tag{2.16}$$

In physical terms, λ is proportional to the tension exercised by the string on the mass.

So far we have shown how to make the Lagrangian formulation of classical mechanics for constrained and non-constrained systems. The method can be applied even for non-inertial frames. Take a mass in two-dimensional free space with the Lagrangian

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2). \tag{2.17}$$

Another observer sees the motion of this free particle in a rotating frame with angular velocity $\boldsymbol{\omega} = (0, 0, \omega)$. New coordinates are related to the old by

$$\begin{aligned} x &= x'\cos\omega t - y'\sin\omega t, \\ y &= y'\cos\omega t + x'\sin\omega t. \end{aligned} \tag{2.18}$$

Velocities are related by

$$\begin{aligned} \dot{x} &= \dot{x}'\cos\omega t - \dot{y}'\sin\omega t - \omega(x'\sin\omega t + y'\cos\omega t) \\ &= (\dot{x}' - \omega y')\cos\omega t - (\dot{y}' + \omega x')\sin\omega t. \\ \dot{y} &= \dot{y}'\cos\omega t + \dot{x}'\sin\omega t + \omega(-y'\sin\omega t + x'\cos\omega t) \\ &= (\dot{y}' + \omega x')\cos\omega t + (\dot{x}' - \omega y')\sin\omega t. \end{aligned} \tag{2.19}$$

Lagrangian in the new frame reads

$$L = \frac{m}{2}\left[(\dot{x}' - \omega y')^2 + (\dot{y}' + \omega x')^2\right]. \tag{2.20}$$

Doing away with the annoying primes, we can also write the Lagrangian in nice vector notation:

$$L = \frac{m}{2}(\dot{\mathbf{r}} + \boldsymbol{\omega}\times\mathbf{r})^2. \tag{2.21}$$

On expanding the square, it appears as though the particle is subject to a velocity-dependent potential $V(\mathbf{r},\dot{\mathbf{r}}) = -(m\omega^2/2)\mathbf{r}^2 - m\dot{\mathbf{r}}\cdot(\boldsymbol{\omega}\times\mathbf{r})$. From

$$\begin{aligned} \frac{\partial L}{\partial \mathbf{r}} &= m(\dot{\mathbf{r}}\times\boldsymbol{\omega} - \boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r})) \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\mathbf{r}}}\right) &= m(\ddot{\mathbf{r}} + \boldsymbol{\omega}\times\dot{\mathbf{r}}), \end{aligned} \tag{2.22}$$

Euler-Lagrange equation follows as

$$\ddot{\mathbf{r}} = -\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) + 2\dot{\mathbf{r}}\times\boldsymbol{\omega}. \tag{2.23}$$

The first term on the r.h.s. is the centrifugal force, the second is the Coriolis force. They are both "apparent" forces caused by the non-inertial nature of the frame of reference.

For rotation about the $\hat{z}$ axis we may write $\boldsymbol{\omega} = \omega\hat{z}$, so that the centrifugal force takes on

$$-\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\mathbf{r}) = +\omega^2\mathbf{r}. \tag{2.24}$$

It is trying to push the particle outwards. A particle at rest in an inertial frame appears to be doing a circular motion

$$\mathbf{r} = R(\cos\omega t, -\sin\omega t) \tag{2.25}$$

in a frame rotating with the angular velocity $\boldsymbol{\omega} = +\omega\hat{z}$. Such motion gives rise to the equation of motion $\ddot{\mathbf{r}} = -\omega^2\mathbf{r}$ while the centrifugal force contributes $+\omega^2\mathbf{r}$ to the acceleration. The discrepancy, equal to $-2\omega^2\mathbf{r}$, is provided by the Coriolis term which gives $2\dot{\mathbf{r}}\times\boldsymbol{\omega} = -2\omega^2\mathbf{r}$ in the case of constant circular

motion. Furthermore, the Coriolis force resembles the Lorentz force acting on a charged particle by the magnetic field. Mathematically, they are the same kind of forces, linearly dependent on the particle's velocity and orthogonal to it.

Now that the subject of velocity-dependent force has been brought up, let's discuss how to write down the Lagrangian for such forces when they are "really there". Specifically we will try to write down the Lagrangian for the motion of a charged particle in electric and magnetic fields, and confirm that the corresponding Lagrange's equation is none other than Lorentz equation. The Lagrangian reads

$$L = \frac{m}{2}\dot{\mathbf{r}}^2 - e(A_0 - \dot{\mathbf{r}} \cdot \mathbf{A}), \tag{2.26}$$

where A_0 and $\mathbf{A}$ are the scalar and vector potentials, respectively. To obtain Lagrange's equation,

$$\frac{\partial L}{\partial \mathbf{r}} = -e\boldsymbol{\nabla} A_0 + e\boldsymbol{\nabla}(\dot{\mathbf{r}} \cdot \mathbf{A}) = \frac{d}{dt}(m\dot{\mathbf{r}} + e\mathbf{A}) = m\ddot{\mathbf{r}} + e\dot{\mathbf{A}}. \tag{2.27}$$

Electric ($\mathbf{E}$) and magnetic ($\mathbf{B}$) fields are obtained from the vector potentials according to

$$\mathbf{E} = -\boldsymbol{\nabla} A_0 - \frac{\partial \mathbf{A}}{\partial t},\ \mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}. \tag{2.28}$$

To cast the vector potentials as electric and magnetic fields, first express

$$\dot{\mathbf{A}} = \frac{\partial \mathbf{A}}{\partial t} + (\dot{\mathbf{r}} \cdot \boldsymbol{\nabla})\mathbf{A}. \tag{2.29}$$

Rearranging some terms,

$$m\ddot{\mathbf{r}} = e\mathbf{E} + e\boldsymbol{\nabla}(\dot{\mathbf{r}} \cdot \mathbf{A}) - e(\dot{\mathbf{r}} \cdot \boldsymbol{\nabla})\mathbf{A}. \tag{2.30}$$

To recognize that the final two terms on the r.h.s. in fact equal the Lorentz force,

$$\partial_i(\dot{x}_j A_j) - \dot{x}_j \partial_j A_i = \dot{x}_j(\partial_i A_j - \partial_j A_i) = \dot{x}_j \varepsilon_{ijk} B_k = (\mathbf{v} \times \mathbf{B})_i. \tag{2.31}$$

We have derived the equation of motion of a charged particle (with charge e) in the presence of electric and magnetic fields, as well as the Lagrangian which produces it.

2.2 Hamiltonian

The Hamiltonian approach is yet another way to reproduce the Newton's force law. Like the Lagrangian formulation, it can be easily generalized to study quantum-mechanical laws as well as the laws of classical physics. As a physicist, one is required to be familiar with both approaches. Actually Hamiltonian

formulation builds upon the Lagrangian formulation in that it requires the definition of a canonical momentum p_i as the conjugate variable of the canonical coordinate q_i, given by the partial differential

$$p_i = \frac{\partial L}{\partial \dot{q}_i}. \tag{2.32}$$

Instead of working with canonical coordinates q_i and its time derivatives, we will be working with q_i and their conjugate momentum p_i. Once p_i is defined in the above manner, we can introduce the quantity, Hamiltonian, as

$$H[q_i, p_i] = p_i \dot{q}_i - L[q_i, \dot{q}_i]. \tag{2.33}$$

Summation over the repeated index i is assumed. Each $\dot{q}_i$ is to be viewed as some combination of p_i's and q_i's. Let's take the derivative of H with respect to q_i and p_i and see what we get.

$$\frac{\partial H}{\partial q_i} = p_j \frac{\partial \dot{q}_j}{\partial q_i} - \frac{\partial L}{\partial q_i} - \frac{\partial \dot{q}_j}{\partial q_i}\frac{\partial L}{\partial \dot{q}_j} = \left(p_j - \frac{\partial L}{\partial \dot{q}_j}\right)\frac{\partial \dot{q}_j}{\partial q_i} - \frac{\partial L}{\partial q_i}. \tag{2.34}$$

The terms inside the parenthesis become zero due to the definition of p_j, while the last term according to the Euler-Lagrange equation equals the time derivative of p_i, hence $\partial H/\partial q_i$ simply equals $-\dot{p}_i$. Differentiation for p_i gives

$$\frac{\partial H}{\partial p_i} = \dot{q}_i + p_j \frac{\partial \dot{q}_j}{\partial p_i} - \frac{\partial L}{\partial \dot{q}_j}\frac{\partial \dot{q}_j}{\partial p_i} = \dot{q}_i + \left(p_j - \frac{\partial L}{\partial \dot{q}_j}\right)\frac{\partial \dot{q}_j}{\partial p_i} = \dot{q}_i. \tag{2.35}$$

In Hamilton's theory, Euler-Lagrange equations are replaced by a pair of equations that read

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}. \tag{2.36}$$

Except for the relative minus sign, p_i and q_i appear in a symmetrical fashion.

Let's apply the Lagrangian $\rightarrow$ Hamiltonian transformation rule just worked out to a simple problem. Starting from the Lagrangian $L = (m/2)\dot{\mathbf{r}}^2 - V(\mathbf{r})$, Hamiltonian follows by first writing down the canonical momentum, $\mathbf{p} = \partial L/\partial \dot{\mathbf{r}} = m\dot{\mathbf{r}}$, and then using it to write $\dot{\mathbf{q}} = \mathbf{p}/m$,

$$H = \mathbf{p} \cdot \left(\frac{\mathbf{p}}{m}\right) - \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}) = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}). \tag{2.37}$$

For a physical trajectory, Hamiltonian is nothing other than the total energy E.

The observation that H plays the role of the total energy can be generalized to the case in which the Lagrangian is given in terms of a generalized set of coordinates q_i and velocities $\dot{q}_i$. With an assumption that the kinetic energy $K(q_i, \dot{q}_i)$ is the only place where the dependence on the velocity $\dot{q}_i$ occurs, we can conclude

$$p_i \equiv \frac{\partial L}{\partial \dot{q}_i} = \frac{\partial K}{\partial \dot{q}_i}. \tag{2.38}$$

Furthermore, we will assume that K is some quadratic function of the velocities, given by $K = \sum_{i \le j} K_{ij} \dot{q}_i \dot{q}_j$. It is straightforward to show that for arbitrary matrix of coefficients K_{ij} one has $\sum_i \dot{q}_i \partial K / \partial \dot{q}_i = 2K$, which gives

$$H = \dot{q}_i p_i - L = 2K - (K - V) = K + V. \tag{2.39}$$

The final expression equals the total energy. The proof does not work for cases of velocity-dependent potential as occurs in electrodynamics.

From the single-particle Hamiltonian (2.37) we derive the Hamilton's equations

$$\dot{\mathbf{r}} = \frac{\partial H}{\partial \mathbf{p}} = \frac{\mathbf{p}}{m}, \quad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{r}} = -\boldsymbol{\nabla} V(\mathbf{r}). \tag{2.40}$$

The first equation tells us that the momentum $\mathbf{p}$ is given by the familiar product of mass and velocity: $m\dot{\mathbf{r}}$. The second equation is the Newton's second law.

Equation (2.36) tells us how the two canonical variables p_i and q_i evolve over time. All other dynamical variables are in turn functions of p_i's and q_i's. Denote such a dynamical variable as $f(p_i, q_i, t)$ where an explicit time dependence is allowed as well. The time evolution of f is governed by the equation

$$\begin{aligned} \frac{df}{dt} &= \frac{\partial f}{\partial t} + \sum_i \left(\frac{\partial f}{\partial p_i} \dot{p}_i + \frac{\partial f}{\partial q_i} \dot{q}_i \right) \\ &= \frac{\partial f}{\partial t} + \sum_i \left(-\frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} + \frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} \right). \end{aligned} \tag{2.41}$$

The final expression in the second line has a neat structure known as the Poisson bracket, which is defined between a pair of variables f and g as

$$[f, g] = \sum_i \left(\frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i} - \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} \right). \tag{2.42}$$

Using the Poisson bracket we can write the time evolution of the variable f as

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + [H, f]. \tag{2.43}$$

This is the generalization of the Hamilton's equations for p_i and q_i first found in Eq. (2.36). In particular Poisson brackets for the generalized coordinates and momenta are

$$[q_i, q_j] = 0 = [p_i, p_j], \quad [p_i, q_j] = \delta_{ij}. \tag{2.44}$$

In the previous section we introduce the Lagrangian of a particle subject to an electromagnetic potential

$$L = \frac{m}{2}\dot{\mathbf{r}}^2 - e(A_0 - \dot{\mathbf{r}} \cdot \mathbf{A}). \tag{2.45}$$

Here the canonical momentum follows $\mathbf{p} = m\dot{\mathbf{r}} + e\mathbf{A}$. It is no longer simply $m\dot{\mathbf{r}}$. Taking $m\dot{\mathbf{r}} = \mathbf{p} - e\mathbf{A}$ everywhere, we get

$$\begin{aligned} H &= \mathbf{p} \cdot \frac{\mathbf{p} - e\mathbf{A}}{m} - \frac{(\mathbf{p} - e\mathbf{A})^2}{2m} + eA_0 - e\mathbf{A} \cdot \frac{\mathbf{p} - e\mathbf{A}}{m} \\ &= \frac{1}{2m}(\mathbf{p} - e\mathbf{A})^2 + eA_0. \end{aligned} \tag{2.46}$$

When the magnetic field is uniform $\mathbf{B} = (0, 0, B)$, the vector potential can be chosen as $\mathbf{A} = (-By, 0, 0)$ giving us the Hamiltonian

$$H = \frac{1}{2m}(p_x + eBy)^2 + \frac{p_y^2}{2m}. \tag{2.47}$$

There are four, first-order equations following from this Hamiltonian. Two of them are

$$\dot{p}_x = 0, \quad \dot{p}_y = -\frac{eB}{m}(p_x + eBy). \tag{2.48}$$

We write the constant p_x as $-eBy_0$, and introduce a Larmor frequency $\omega_L = eB/m$. Then

$$\dot{p}_y = -m\omega_L^2(y - y_0). \tag{2.49}$$

The other two equations are

$$\dot{x} = \frac{1}{m}(p_x + eBy) = \omega_L(y - y_0), \quad \dot{y} = \frac{p_y}{m}. \tag{2.50}$$

By comparing the above two equations we also learn $\dot{p}_y + m\omega_L\dot{x} = 0$, which means $p_y = m\omega_L(x_0 - x)$. The only two independent equations left to solve can be chosen as

$$\dot{x} = \omega_L(y - y_0), \quad \dot{y} = -\omega_L(x - x_0). \tag{2.51}$$

They are solved as $x - x_0 = A\cos(\omega_L t + \alpha)$, $y - y_0 = A\sin(\omega_L t + \alpha)$. In quantum mechanics, the Larmor frequency becomes known as the cyclotron frequency. The Landau level spacing becomes $\hbar$ times the cyclotron frequency.

The same problem can be solved a little more formally by use of the Poisson brackets. Defining $\mathbf{\Pi} = \mathbf{p} + e\mathbf{A}$, one can re-write the Hamiltonian

$$H = \frac{1}{2m}(\Pi_x^2 + \Pi_y^2). \tag{2.52}$$

First one can show that the Poisson bracket for Π_x and Π_y is

$$[\Pi_x, \Pi_y] = -eB. \tag{2.53}$$

Next one works out the equation of motion for Π_x and Π_y,

$$\begin{aligned} \dot{\Pi}_x &= [H, \Pi_x] = \frac{eB}{m}\Pi_y, \\ \dot{\Pi}_y &= [H, \Pi_y] = -\frac{eB}{m}\Pi_x. \end{aligned} \tag{2.54}$$

The motion is that of a circular motion with the angular frequency $\omega = eB/m$.

2.3 Problems

1. Prove the energy conservation law for multi-particle systems governed by

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i, \tag{2.55}$$

where $i = 1, \cdots, N$ runs over all the particles each carrying a mass m_i. The force $\mathbf{F}_i$ acting on the i-th particle is derived from the pair-wise interaction energy $V = \sum_{i<j} V(\mathbf{r}_i - \mathbf{r}_j)$.

2. Take the action of a one-dimensional free particle $\mathcal{L} = \int_{t_1}^{t_2} (m/2)\dot{x}^2 dt$. We know the physical motion is that of constant velocity

$$x_p(t) = L\frac{t - t_1}{t_2 - t_1} \tag{2.56}$$

consistent with boundary conditions $x(t_1) = 0, x(t_2) = L$. Suppose an alternative path to the physical one had been chosen,

$$x(t) = x_p(t) + \eta \sin\left(\pi \frac{t - t_1}{t_2 - t_1}\right). \tag{2.57}$$

Prove that the new path $x(t)$ gives more action than the physical path $x_p(t)$.

3. A pendulum of length l and mass m has its point of support attached to a block of mass M that is free to slide horizontally. Coordinates of each mass are given by $(X, 0)$ and (x, y), with the constraint $(x-X)^2+y^2 = l^2$. Write down the Lagrangian for this system and derive the Euler-Lagrange equation for X and θ, the angle of the pendulum swing.

4. A particle of mass m moves in the plane subject to a central force field with the potential energy $V(r)$. Write down the corresponding Lagrangian in the cylindrical coordinates and derive the Euler-Lagrange equations of motion.

5. Prove Eq. (2.22).

6. Based on Eq. (2.23), what can an observer in a rotating frame conclude about the conservation of kinetic energy $E = (m/2)\dot{\mathbf{r}}^2$? Is it conserved? If not, what is the time dependence of the energy dE/dt?

7. Derive the equation of motion starting from the Lagrangian for the relativistic particle $L = -mc^2\sqrt{1 - \dot{\mathbf{r}} \cdot \dot{\mathbf{r}}/c^2} - V(\mathbf{r})$. Show that it reduces to Newton's law in the limit $|\dot{\mathbf{r}}| \ll c$.

8. Prove Eq. (2.44).

9. A spherical pendulum of length l has a mass m attached at its end which is free to swing in three dimensions. Its position is described by two angles of the spherical coordinates (θ, ϕ). Write down the Lagrangian and derive the equation of motion. There is a constant force acting vertically, $\mathbf{F} = -mg\hat{z}$. We will get back to the analysis of the derived equation of motion in a later chapter.

10. Derive the Hamiltonian of a simple harmonic oscillator starting from the Lagrangian $L = (m/2)\dot{x}^2 - (k/2)x^2$.

11. Starting from the Lagrangian for the particle of mass m in a central force field $V(r)$ in two dimensions, derive the corresponding Hamiltonian H, then derive Hamilton's equation of motion.

12. Starting from the Lagrangian (2.21) in the rotating frame, derive the Hamiltonian for the free particle in the rotating frame of constant angular velocity $\boldsymbol{\omega}$.

13. Angular momentum vector $\mathbf{L} = (L_x, L_y, L_z)$ is defined as $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. Derive the Poisson bracket relation among the components: $[L_x, L_y] = -L_z$ and its cyclic permutations. Further prove that $[\mathbf{L}^2, L_i] = 0$ for $i = x, y, z$. Both relations have their analogues in quantum-mechanical commutators.

14. Show that any function of the radius $r = |\mathbf{r}|$, $\phi(r)$, has a Poisson bracket with L_z equal to zero: $[\phi(r), L_z] = 0$.

Chapter 3

Symmetries and Conservation Laws

3.1 View from Newtonian Mechanics

A particle in motion subject to a conservative force, $\mathbf{F} = -\boldsymbol{\nabla} V$ has its total energy, $(m/2)\dot{\mathbf{r}}^2 + V$ conserved throughout the motion. We showed this in the previous chapter. In turn, it says that any non-conservative system does not have the luxury of maintaining the constant energy. The simplest way to introduce non-conservative forces is to allow explicit time dependence in the force. For instance, take the gravitational force exerted by one object on another, $V(r) = -k/r$. Suppose the source of gravitational force at the origin is shaken in a periodic manner by some external means. The relative position vector $\mathbf{r}$ will also pick up the oscillatory dependence and becomes $\mathbf{r} - \mathbf{r}_0(t)$. The potential $V(\mathbf{r} - \mathbf{r}_0(t))$ is no longer energy-conserving even though one still has

$$m\ddot{\mathbf{r}} = -\boldsymbol{\nabla} V(\mathbf{r} - \mathbf{r}_0(t)). \tag{3.1}$$

To see why, multiply both sides by $\dot{\mathbf{r}}$,

$$\frac{d}{dt}\left(\frac{m}{2}\dot{\mathbf{r}}^2\right) = -\frac{d\mathbf{r}}{dt} \cdot \boldsymbol{\nabla} V(\mathbf{r} - \mathbf{r}_0(t)) = -\frac{dV}{dt} + \frac{\partial V}{\partial t}. \tag{3.2}$$

Now the total energy $E = K + V$ obeys

$$\frac{dE}{dt} = \frac{\partial V}{\partial t} = -\dot{\mathbf{r}}_0 \cdot \boldsymbol{\nabla} V, \tag{3.3}$$

which clearly shows the lack of energy conservation for time-dependent potential. The examples shows that energy conservation is closely tied to the time(t)-independence of the system, or the isotropy of time.

Angular momentum of a particle is given by $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. Whether this quantity is conserved or not depends on

$$\frac{d\mathbf{L}}{dt} = \mathbf{r} \times \dot{\mathbf{p}} = -\mathbf{r} \times \boldsymbol{\nabla} V. \tag{3.4}$$

Generally, the r.h.s. (torque) would be non-zero, but in the case of central potential $V = V(r)$, the gradient becomes $\boldsymbol{\nabla} V = \hat{r} V'(r)$, and $\mathbf{r} \times \boldsymbol{\nabla} V = 0$. For the central potential acting on the particle, $\mathbf{L}$ is conserved. Central potential $V(r)$ depends only on the distance from the origin, making it invariant under an arbitrary orthogonal rotation. We can infer that angular momentum conservation has to do with the isotropy of space.

We can also talk about "partial" angular momentum conservation. If the potential is given in the form $V(r, z)$, $r = (x^2+y^2)^{1/2}$, we find $\boldsymbol{\nabla} V = \hat{z}\frac{\partial V}{\partial z} + \hat{r}\frac{\partial V}{\partial r}$. For the z-component of $\mathbf{L}$,

$$\frac{dL_z}{dt} = xF_y - yF_x = 0. \tag{3.5}$$

The conserved component is related to the axis about which rotation leaves potential form preserved.

Linear momentum conservation rests on

$$\frac{d\mathbf{p}}{dt} = -\boldsymbol{\nabla} V. \tag{3.6}$$

It is conserved when $V = V_0$, for a free particle. As in the angular momentum case, we can talk about "partial" momentum conservation. If $V(x)$ is only a function of x, p_y and p_z are conserved. The conserved directions are those along which the potential function remains invariant.

3.2 View from Lagrangian Mechanics

The case of time-independent force corresponds to the Lagrangian L which does not contain t as an explicit variable. In such a case the total time derivative of the Lagrangian L is

$$\begin{aligned} \frac{dL}{dt} &= \frac{\partial L}{\partial q_i}\dot{q}_i + \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i \\ &= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\dot{q}_i + \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i = \frac{d}{dt}\left(\dot{q}_i \frac{\partial L}{\partial \dot{q}_i}\right), \end{aligned} \tag{3.7}$$

the second line following from the Euler-Lagrange equation itself. Taking the difference of the two sides, we find an invariant of time

$$\frac{d}{dt}\left(\dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L\right) = 0. \tag{3.8}$$

In other words, the quantity we previously defined as the Hamiltonian H, when viewed as a function of the generalized coordinates q_i and generalized velocities $\dot{q}_i$, does not change over time. For physically allows paths (obeying the Euler-Lagrange equation), H is a constant of motion provided $\partial L/\partial t = 0$. We know very well that H is the total energy.

In general, if some components of the generalized coordinates were missing from the Lagrangian such that $\partial L/\partial q_i = 0$, there is a corresponding conserved quantity $p_i = \partial L/\partial \dot{q}_i$.

$$\dot{p}_i = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{\partial L}{\partial q_i} = 0. \tag{3.9}$$

This is the generalization of the notion of "partial" linear and angular momentum conservation discussed in the previous section. From symmetry perspective the absence of a particular coordinate q_i means that the physical system remains invariant under the displacement of that coordinates by an arbitrary amount $q_i \to q_i + a_i$. If q_i were a rectilinear coordinate the invariance is a consequence of the "homogeneity" of space along that direction. If q_i were an angular coordinates the invariance is a consequence of the "isotropy" of space about the axis associated with the angle.

As a simple application of the above-mentioned principle relating invariance of the Lagrangian to a conservation law, consider a homogeneous system consisting of N-particles whose Lagrangian is

$$L = \frac{1}{2}\sum_i m_i \dot{\mathbf{r}}_i^2 - \sum_{i<j} V(|\mathbf{r}_i - \mathbf{r}_j|). \tag{3.10}$$

The central force between a pair of particles is mediated by the potential $V(|\mathbf{r}_i - \mathbf{r}_j|)$. This Lagrangian is invariant under the translation of all the coordinates by a uniform amount $\mathbf{r}_i \to \mathbf{r}_i + \mathbf{u}$. From the invariance,

$$L[\mathbf{r}_i + \mathbf{u}, \dot{\mathbf{r}}_i] = L[\mathbf{r}_i, \dot{\mathbf{r}}_i] \tag{3.11}$$

follows the conclusion

$$\mathbf{u} \cdot \sum_i \frac{\partial L}{\partial \mathbf{r}_i} = \mathbf{u} \cdot \sum_i \dot{\mathbf{p}}_i = \mathbf{u} \cdot \frac{d}{dt}\left(\sum_i \mathbf{p}_i\right) = 0. \tag{3.12}$$

The relation has to hold for arbitrary $\mathbf{u}$, hence the total linear momentum $\mathbf{P} = \sum_i \mathbf{p}_i$ must be conserved.

When specialized to the two-body problem we also conclude

$$\sum_{i=1,2} \frac{\partial L}{\partial \mathbf{r}_i} = -\sum_{i=1,2} \frac{\partial V}{\partial \mathbf{r}_i} = \sum_{i=1,2} \mathbf{F}_i = 0. \tag{3.13}$$

That is, the force acting on 1 by 2 ($\mathbf{F}_1$) must be equal and opposite to the force acting on 2 by 1 ($\mathbf{F}_2$). The third law of Newton, action-reaction principle, follows naturally from the momentum conservation of the two-body problem, which in turn follows from the homogeneity of the Lagrangian under the uniform translation.

If a system is invariant under rotation, the Lagrangian keeps the same form under

$$\mathbf{r}_i \to \mathbf{r}_i + s\hat{\boldsymbol{n}} \times \mathbf{r}_i,\ \dot{\mathbf{r}}_i \to \dot{\mathbf{r}}_i + s\hat{\boldsymbol{n}} \times \dot{\mathbf{r}}_i, \tag{3.14}$$

where $\hat{\boldsymbol{n}}$ serves as the rotation axis and s is the (small) amount of rotation about the axis. It follows that

$$\begin{aligned}\sum_i \left(\hat{\boldsymbol{n}} \times \mathbf{r}_i \cdot \frac{\partial L}{\partial \mathbf{r}_i} + \hat{\boldsymbol{n}} \times \dot{\mathbf{r}}_i \cdot \frac{\partial L}{\partial \dot{\mathbf{r}}_i}\right) &= \hat{\boldsymbol{n}} \cdot \sum_i \left(\mathbf{r}_i \times \dot{\mathbf{p}}_i + \dot{\mathbf{r}}_i \times \mathbf{p}_i\right) \\ = \hat{\boldsymbol{n}} \cdot \frac{d}{dt}\left(\sum_i \mathbf{r}_i \times \mathbf{p}_i\right) &= \frac{d}{dt}\left(\hat{\boldsymbol{n}} \cdot \mathbf{L}\right) = 0.\end{aligned} \tag{3.15}$$

The total angular momentum $\mathbf{L}$ projected onto the direction of rotation $\hat{\boldsymbol{n}}$ is thus conserved.

The two examples discussed here, one of conservation of total linear momentum and the other one of total angular momentum, are instances of what is known as Noether's theorem. The theorem states that if a Lagrangian remains invariant under a change change of the coordinates parameterized by the one constant α, then there exists one conserved quantity associated with such a continuous change. It also tells us that conservation laws are associated with the continuous symmetries of the Lagrangian.

3.3 View from Hamiltonian Mechanics

For time-independent Lagrangian, we saw that the energy, i.e. the Hamiltonian H is independent of time, so that $dH/dt = 0$. To see this directly in the Hamiltonian formulation, calculate

$$\frac{dH}{dt} = \sum_i \frac{\partial H}{\partial q_i}\dot{q}_i + \sum_i \frac{\partial H}{\partial p_i}\dot{p}_i = -\dot{p}_i\dot{q}_i + \dot{q}_i\dot{p}_i = 0. \tag{3.16}$$

Take some quantity $A(p_i, q_i)$ which depends on the coordinates q_i and their canonical momenta p_i. To see whether the quantity is conserved, one should calculate

$$\frac{dA}{dt} = \frac{\partial A}{\partial q_i}\dot{q}_i + \frac{\partial A}{\partial p_i}\dot{p}_i = \frac{\partial A}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial H}{\partial q_i} = [H, A]. \tag{3.17}$$

This quantity is the Poisson bracket already introduced in the previous chapter. If a physical quantity A is a constant motion, it must commute with the Hamiltonian H in the sense that its Poisson bracket with H is zero. As a trivial consequence, any observable A which is a function of H and not of the p_i and q_i individually is a constant of motion.

$$\begin{aligned}\frac{dA}{dt} &= \frac{\partial A}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial A}{\partial p_i}\frac{\partial H}{\partial q_i} = \frac{\partial A}{\partial H}\frac{\partial H}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial A}{\partial H}\frac{\partial H}{\partial p_i}\frac{\partial H}{\partial q_i} \\ &= \frac{\partial A}{\partial H}[H, H] = 0.\end{aligned} \tag{3.18}$$

A famous example of such an observable quantity is the Botlzmann distribution function $\rho = \exp(-H/k_B T)$.

Conservation of linear momentum would follow from

$$\frac{d\mathbf{p}}{dt} = [H, \mathbf{p}]. \tag{3.19}$$

If the Hamiltonian is that of a free particle, $H = \mathbf{p}^2/2m$, it follows immediately that $[H, \mathbf{p}] = 0$, hence $\dot{\mathbf{p}} = 0$. Conservation of the angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ for the Hamiltonian $H = \mathbf{p}^2/2m + V(|\mathbf{r}|)$ can be shown using the similar Poisson bracket algebra.

3.4 Problems

1. Prove the conservation of angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ for the Hamiltonian $H = \mathbf{p}^2/2m + V(|\mathbf{r}|)$ by working out the Poisson bracket $[H, \mathbf{L}]$.

2. A particle moving in the $x < 0$ region of space with potential energy $V = V_1$ is incident on the $x = 0$ plane at an angle θ_1 with the normal and emerges on the $x > 0$ space at an angle θ_2 where the potential is $V = V_2$. Calculate the ratio of the incident and emergent angles by calculating $\sin\theta_1 / \sin\theta_2$. The answer must be expressed in terms of the potential energy difference $V_1 - V_2$, and the kinetic energy of the incident particle $m\mathbf{v}_1^2/2$.

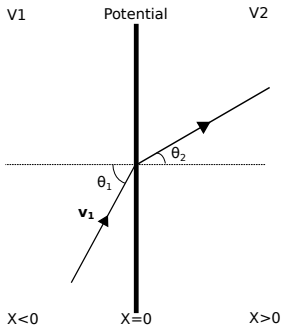

Figure 3.1: Problem 3.2

Chapter 4

One-dimensional Motion

4.1 First-integral Approach

One-dimensional motion in classical mechanics refers to a set of situations wherein the dynamics can be reduced to that of a single variable. All other variables can be "eliminated" by use of conservation laws.

A prototypical Lagrangian of a one-dimensional motion is

$$L = \frac{1}{2}m\dot{x}^2 - V(x). \tag{4.1}$$

Conservation of energy ensures that $E = \frac{1}{2}m\dot{x}^2 + V(x)$ will remain constant throughout the motion. We infer that

$$\frac{dx}{dt} = \pm\sqrt{E - V(x)}. \tag{4.2}$$

For convenience take the positive branch, and re-write the first-order differential equation in the form

$$dt = \frac{dx}{\sqrt{E - V(x)}}. \tag{4.3}$$

Upon integrating, we obtain

$$t_2 - t_1 = \sqrt{\frac{m}{2}}\int_{x_1}^{x_2} \frac{dx}{\sqrt{E - V(x)}}. \tag{4.4}$$

Obviously the motion is only allowed in the region where $V(x) < E$, which in turn defines the turning points $E = V(x)$. In the case when the motion is confined at both ends, with two turning points x_l and x_r, one should be able to derive the period of the motion as

$$T = \sqrt{2m}\int_{x_l}^{x_r} \frac{dx}{\sqrt{E - V(x)}}. \tag{4.5}$$

Applied to the problem of simple harmonic oscillator where we can take $E = (k/2)X^2$, where X is the maximum displacement, we have

$$T = \sqrt{2m}\int_{-X}^{X}\frac{dx}{\sqrt{k/2(X^2-x^2)}} = \frac{2}{\omega_0}\int_{-1}^{1}\frac{dy}{\sqrt{1-y^2}} = \frac{2\pi}{\omega_0}. \tag{4.6}$$

The desired period $T = 2\pi/\omega_0$ is obtained. It is independent of the amplitude of harmonic oscillation X.

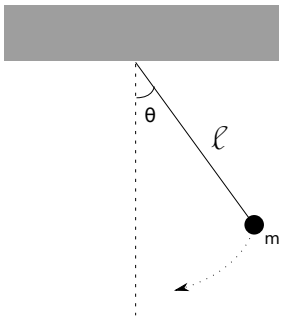

Figure 4.1: A simple pendulum.

With the more general motion of a swinging pendulum we have the total energy given by

$$\frac{1}{2}ml^2\dot{\theta}^2 - mgl\cos\theta = -mgl\cos\theta_0. \tag{4.7}$$

We used θ_0 to indicate the maximum angle of swing. For the angular velocity $\dot{\theta}$ we get

$$\begin{aligned}&\frac{d\theta}{dt} = \sqrt{\frac{2g}{l}(\cos\theta - \cos\theta_0)}\\ &\Rightarrow t_2 - t_1 = \frac{1}{\sqrt{2}\omega_0}\int_{\theta_1}^{\theta_2}\frac{d\theta}{\sqrt{\cos\theta-\cos\theta_0}},\ \omega_0 = \sqrt{\frac{g}{l}}.\end{aligned} \tag{4.8}$$

For small oscillations we may approximate the cosines by their Taylor expansion,

$$t_2 - t_1 \approx \frac{1}{\omega_0}\int_{\theta_1}^{\theta_2}\frac{d\theta}{\sqrt{\theta_0^2-\theta^2}}, \tag{4.9}$$

which would reproduce the harmonic oscillator result. For large swings the Taylor expansion no longer works, and we have to evaluate the integral numerically. The period of a large oscillation is given by

$$\omega_0 T = 2\sqrt{2}\int_0^{\theta_0}\frac{d\theta}{\sqrt{\cos\theta-\cos\theta_0}}. \tag{4.10}$$

Because the integral is impossible to evaluate exactly, it is instead given a grand name "Jacobi complete elliptic integral of the first kind". For small angle θ_0 the period is given by

$$\omega_0 T = 2\pi \left(1 + \frac{1}{16}\theta_0^2 + \cdots\right). \tag{4.11}$$

Even for a fairly large angle of swing the first correction term proves quite small compared to 1, and presumably the second correction term of order θ_0^4 even smaller. In practical calculations retaining the first correction will be adequate.

4.2 Effective One-Dimensional Motion

A system of two point particles interacting by the central force can always be reduced to the single-particle motion. Starting from

$$L = \frac{1}{2}m_1\dot{\mathbf{r}}_1^2 + \frac{1}{2}m_2\dot{\mathbf{r}}_2^2 - V(|\mathbf{r}_1 - \mathbf{r}_2|), \tag{4.12}$$

we introduce the reduced mass m and the relative coordinates $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$. In this case, the Lagrangian can be re-written as

$$L = \frac{1}{2}M\dot{\mathbf{R}}^2 + \frac{1}{2}m\dot{\mathbf{r}}^2 - V(r), \tag{4.13}$$

$M = m_1 + m_2$, $M\dot{\mathbf{R}} = m_1\dot{\mathbf{r}}_1 + m_2\dot{\mathbf{r}}_2$. The total momentum is a constant of motion, and can be dropped from the Lagrangian. The relevant problem is that of the relative motion described by $\mathbf{r}$.

$$L = \frac{1}{2}m\dot{\mathbf{r}}^2 - V(r). \tag{4.14}$$

As discussed in the previous chapter, this Lagrangian also happens to have the conserved angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, $\mathbf{p} = m\dot{\mathbf{r}}$. Choosing the conserved $\mathbf{L}$ direction to be the z-direction, it means that the motion will only take place in the xy plane. Employing the cylindrical coordinates, Lagrangian becomes

$$L = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - V(r). \tag{4.15}$$

Lagrange's equations are

$$m\ddot{r} = mr\dot{\theta}^2 - V'(r), \quad \frac{d}{dt}(mr^2\dot{\theta}) = 0. \tag{4.16}$$

The conserved quantity $mr^2\dot{\theta}$ is the angular momentum about the z-axis, L_z. Another way to see the conservation is to note that the coordinate θ does not appear explicitly in the Lagrangian. For all such coordinates, we saw in the previous chapter, there is a corresponding conserved quantity. Denoting the

conserved angular momentum $mr^2\dot{\theta} = L_z$ allows us to re-write the remaining equation as

$$m\ddot{r} = \frac{L_z^2}{mr^3} - V'(r). \tag{4.17}$$

Indeed, the two-body problem interacting by the central force is reduced to that of a single mass executing a one-dimensional motion subject to the effective potential $V_{\rm eff}(r) = V(r)+L_z^2/2mr^2$. Emergence of the effective potential instead of the bare one $V(r)$ is a consequence of "projecting" the problem from higher dimensions to one. Such phenomena, i.e. renormalization of the bare interaction while eliminating the degrees of freedom, occurs in all branches of many-body physics. Solution of the differential equation for $V(r) = -k/r$ will lead to Kepler's elliptic orbits and all that.

Although not quite as powerful as the full solution, one can extract some relevant information about the motion by considering the constant of motion $E = (m/2)(\dot{r}^2 + r^2\dot{\theta}^2) + V(r) = (m/2)\dot{r}^2 + L_z^2/2mr^2 + V(r)$, invoking the constancy of angular momentum to write down the latter. So,

$$\frac{dr}{dt} = \sqrt{\frac{2}{m}[E - V_{\rm eff}(r)]}. \tag{4.18}$$

Substitution of the central potential $V(r) = -k/r$ leads to the picture of the motion for r oscillating between the two extremum radii set by

$$E + \frac{k}{r} - \frac{L_z^2}{2mr^2} = 0 \Rightarrow r^2 + \frac{k}{E}r - \frac{L_z^2}{2mE} = 0. \tag{4.19}$$

We obtain the two roots as

$$r_{1,2} = -\frac{k}{2E} \pm \sqrt{\left(\frac{k}{2E}\right)^2 + \frac{L_z^2}{2mE}}. \tag{4.20}$$

For some choice of $E < 0$ the term inside the square root vanishes:

$$E_c = -\frac{mk^2}{2L_z^2}. \tag{4.21}$$

Only for this particular energy is the motion perfectly circular. For other energies, the oscillation of the radius between r_1 and r_2 results in the elliptic orbits.

Using the constant of motion $mr^2 d\theta/dt = L_z$ one can convert Eq. (4.18) to

$$\frac{dr}{d\theta} = \frac{dr/dt}{d\theta/dt} = \frac{mr^2}{L_z}\frac{dr}{dt} = \frac{r^2}{L_z}\sqrt{2m[E - V_{\rm eff}(r)]}. \tag{4.22}$$

Integration yields the relation

$$\theta - \theta_0 = \int_{r_0}^{r} \frac{L_z}{r'^2\sqrt{2m[E - V_{\text{eff}}(r')]}} dr'. \tag{4.23}$$

In particular the angle picked up during one cycle of $r = r_2$ to $r = r_1$ and back, written $\Delta\theta$,

$$\Delta\theta = 2L_z \int_{r_2}^{r_1} \frac{dr'}{r'^2\sqrt{2m[E - V_{\text{eff}}(r')]}}, \tag{4.24}$$

has to be a rational multiple of 2π if the orbits were to repeat itself after a finite time.

Another seemingly two-dimensional problem reducible to a one-dimensional one is the motion of a spherical pendulum. The cartesian coordinates of a spherical pendulum are given by

$$x = l\sin\theta\cos\phi,\ y = l\sin\theta\sin\phi,\ z = -l\cos\theta. \tag{4.25}$$

The Lagrangian in terms of θ and ϕ reads

$$\begin{aligned} L &= \frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - mgz \\ &= \frac{1}{2}ml^2(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta) + mgl\cos\theta. \end{aligned} \tag{4.26}$$

Note that the coordinate θ does not appear, implying the conservation of the associated momentum $\partial L/\partial\dot{\theta} = ml^2\dot{\phi}\sin^2\theta = L_z$. This is the angular momentum component along the direction which also serves as the axis of rotation for angle ϕ. The equation of motion for θ follows from the Lagrangian

$$ml^2\ddot{\theta} = ml^2\dot{\phi}^2\sin\theta\cos\theta - mgl\sin\theta, \tag{4.27}$$

but $\dot{\phi}$ can be eliminated using the constant of motion just derived, $\dot{\phi} = L_z/ml^2\sin^2\theta$. The whole thing can be summarized as

$$\ddot{\theta} = -\frac{\partial V(\theta)}{\partial\theta} \tag{4.28}$$

where the effective potential $V(\theta)$ is

$$V(\theta) = -\frac{g}{l}\cos\theta + \frac{L_z^2}{2m^2l^4}\frac{1}{\sin^2\theta}. \tag{4.29}$$

It appears to be a one-dimensional problem. The reason for introducing a term new $V(\theta)$ is that it is now rather see to express the energy conservation as

$$E = \frac{1}{2}\dot{\theta}^2 + V(\theta). \tag{4.30}$$

Solving t in terms of θ proceeds as

$$t_2 - t_1 = \frac{1}{\sqrt{2}} \int_{\theta_1}^{\theta_2} \frac{d\theta}{\sqrt{E - V(\theta)}}. \tag{4.31}$$

For θ values corresponding to the minimum of $V(\theta)$ we obtain $\ddot{\theta} = 0$. The spherical pendulum stays at a constant angle θ_0. For values of θ deviating slightly from θ_0 we can expand $V(\theta)$ to quadratic order and obtain a harmonic oscillator equation

$$\ddot{\theta} = -\omega^2 \theta. \tag{4.32}$$

4.3 Problems

1. Prove that Eq. (4.10) is equivalent to $4K(\sin\theta_0/2)$ where

$$K(x) \equiv \int_0^{\pi/2} \frac{dy}{\sqrt{1 - x^2 \sin^2 y}} \tag{4.33}$$

   is known as the complete elliptic integral of the first kind.

2. Evaluate Eq. (4.10) numerically and plot the period as a function of θ_0 ranging from 0 to π. Compare the plot with the approximate result $\omega_0 T \approx 2\pi(1 + \theta_0^2/16)$.

3. Prove Eq. (4.13).

4. Derive the oscillation frequency ω in Eq. (4.32).

Chapter 5

Harmonic Oscillation

5.1 Coupled Harmonic Oscillation

5.1.1 Generality

For a coupled harmonic oscillator the Lagrangian is given by

$$L = \frac{1}{2} T_{ij} \dot{q}_i \dot{q}_j - \frac{1}{2} V_{ij} q_i q_j. \tag{5.1}$$

The sum runs over both indices $1 \leq i, j \leq N$. One can always choose T_{ij} and V_{ij} to be symmetric, $T_{ji} = T_{ij}$, $V_{ji} = V_{ij}$. Lagrange's equations follow as

$$T_{ij} \ddot{q}_j = -V_{ij} q_j \Rightarrow \boldsymbol{T}\ddot{\boldsymbol{q}} = -\boldsymbol{V}\boldsymbol{q}. \tag{5.2}$$

Assume that the matrix $\boldsymbol{T}$ can be inverted, and we would have

$$\ddot{\boldsymbol{q}} = -\boldsymbol{T}^{-1}\boldsymbol{V}\boldsymbol{q} = -\boldsymbol{F}\boldsymbol{q}. \tag{5.3}$$

In a normal mode, all components of $\boldsymbol{q}$ oscillate with the same frequency ω and we may write $\boldsymbol{q}(t) = e^{-i\omega t}\mathbf{q}$ with the time-independent $\mathbf{q}$-vector. Then we arrive at an eigenvalue equation

$$\omega^2 \boldsymbol{q} = \boldsymbol{F}\boldsymbol{q}, \quad \boldsymbol{q} = \begin{pmatrix} q_1 \\ \vdots \\ q_N \end{pmatrix}. \tag{5.4}$$

If $\boldsymbol{F}$ were symmetric, we would know for sure that ω^2 are all real-valued. Even though now $\boldsymbol{F}$ is not guaranteed to be a symmetric matrix, we can still show that ω^2 must be real. To see this, we go back a step and write the eigenvalue problem as

$$\omega^2 \boldsymbol{T}\boldsymbol{q} = \boldsymbol{V}\boldsymbol{q}. \tag{5.5}$$

Taking the inner product on the left with the conjugate of $\boldsymbol{q}$ gives

$$\omega^2 = \frac{\boldsymbol{q}^\dagger \boldsymbol{V} \boldsymbol{q}}{\boldsymbol{q}^\dagger \boldsymbol{T} \boldsymbol{q}}, \quad \boldsymbol{q}^\dagger = (q_1^*, \cdots, q_N^*). \tag{5.6}$$

Due to the fact that both $\boldsymbol{T}$ and $\boldsymbol{V}$ are real and symmetric, it follows that both the denominator and the numerators are real. For instance

$$\left(\sum_{i,j} q_i^* V_{ij} q_j \right)^* = \sum_{i,j} V_{ij} q_i^* q_j = \text{real}. \tag{5.7}$$

Hence ω^2 must be real, too. It is still not guaranteed that ω^2 would be positive, though. Positivivity of ω^2 should follow from the dynamical stability of the Lagrangian itself. To see this point write q_i generally in terms of its real and imaginary parts, $q_i = a_i + ib_i$. Then

$$\sum_{i,j} q_i^* V_{ij} q_j = \sum_{i,j} a_i V_{ij} a_j + \sum_{i,j} b_i V_{ij} b_j. \tag{5.8}$$

Each term on the right side represents the change in the potential energy due to an arbitrary displacement vector $\boldsymbol{a}$ and $\boldsymbol{b}$. Expanding around an stable minimum, by definition, implies that a small arbitrary displacement leads to an *increase* in energy. Hence V_{ij} must be that the two terms on the right are both positive for arbitrary real-valued vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. A similar argument can be made regarding the positivity of $\boldsymbol{q}^\dagger \boldsymbol{T} \boldsymbol{q}$, saying that motion can only increase the kinetic energy and not the other way around.

The structure of the normal mode equation (5.5) is such that one cannot make a distinction between solutions of positive ω and negative ω. For convenience we may always choose $\omega > 0$. For each eigenmode ω_α we have the corresponding eigenvector $\boldsymbol{q}_\alpha$ satisfying

$$\omega_\alpha^2 \boldsymbol{T} \boldsymbol{q}_\alpha = \boldsymbol{V} \boldsymbol{q}_\alpha, \tag{5.9}$$

for $1 \leq \alpha \leq N$. For a linear problem, an arbitrary linear superposition of such solutions is also a solution, hence the most general solution to the coupled oscillator problem is

$$\boldsymbol{q}(t) = \sum_{\alpha=1}^{N} A_\alpha \boldsymbol{q}_\alpha e^{i\omega_\alpha t}, \tag{5.10}$$

with an arbitrary set of complex coefficients A_α. Of course the physical solution is recovered by taking the real part of the above solution.

At this point it is useful to point out the analogy of the normal-mode solutions to the coupled oscillator problem to the quantum mechanical solution. One might regard $\boldsymbol{q}_\alpha$ as an eigenvector representing the α-th eigenstate. Then

Eq. (5.9) has a nice interpretation as the average of the potential energy "operator" $\boldsymbol{V}$ divided by the average of the kinetic energy "operator" $\boldsymbol{T}$ for the n-th eigenstate $|n\rangle$, as formally given by

$$\omega^2 = \frac{\langle n|\boldsymbol{V}|n\rangle}{\langle n|\boldsymbol{T}|n\rangle}. \tag{5.11}$$

Closely related to the normal modes $\boldsymbol{q}_\alpha$ just found above is the notion of normal coordinates, written θ_α, $1 \leq \alpha \leq N$. By definition these are the linear combinations of the original coordinates

$$\theta_\alpha = \sum_i X_{\alpha i} q_i, \quad \dot{\theta}_\alpha = \sum_i X_{\alpha i} \dot{q}_i, \tag{5.12}$$

such that the Lagrangian becomes normal,

$$L = \frac{1}{2} T_{ij} \dot{q}_i \dot{q}_j - \frac{1}{2} V_{ij} q_i q_j = \sum_\alpha \frac{1}{2} m_\alpha \left(\dot{\theta}_\alpha^2 - \omega_\alpha^2 \theta_\alpha^2 \right). \tag{5.13}$$

In effect we are requiring that the two conditions be met simultaneously:

$$\begin{aligned} T_{ij} \dot{q}_i \dot{q}_j &= \sum_\alpha m_\alpha \dot{\theta}_\alpha^2 = \left(\sum_\alpha m_\alpha X_{\alpha i} X_{\alpha j} \right) q_i q_j \\ V_{ij} q_i q_j &= \sum_\alpha m_\alpha \omega_\alpha^2 \theta_\alpha^2 = \left(\sum_\alpha m_\alpha \omega_\alpha^2 X_{\alpha i} X_{\alpha j} \right) \dot{q}_i \dot{q}_j. \end{aligned} \tag{5.14}$$

The two relations can be summarized compactly as matrix relations,

$$\boldsymbol{T} = \boldsymbol{X}^T \boldsymbol{M} \boldsymbol{X}, \quad \boldsymbol{V} = \boldsymbol{X}^T \boldsymbol{M} \boldsymbol{\Omega}^2 \boldsymbol{X}. \tag{5.15}$$

Two diagonal matrices $\boldsymbol{M} = \mathrm{diag}(m_1, \cdots, m_N)$, and $\boldsymbol{\Omega}^2 = (\omega_1^2, \cdots, \omega_N^2)$, have been defined. On the other hand, the normal mode equations (5.9) for all the $1 \leq \alpha \leq N$ can be summarized as a matrix equation:

$$\boldsymbol{T} \boldsymbol{Q} \boldsymbol{\Omega}^2 = \boldsymbol{V} \boldsymbol{Q}, \tag{5.16}$$

where $\boldsymbol{Q} = \left(\boldsymbol{q}_1 \cdots \boldsymbol{q}_N \right)$ is a collection of normal mode solutions arranged as a matrix. Inserting Eq. (5.15) into Eq. (5.16) gives the relation

$$\boldsymbol{X}^T \boldsymbol{M} \boldsymbol{X} \boldsymbol{Q} \boldsymbol{\Omega}^2 = \boldsymbol{X}^T \boldsymbol{M} \boldsymbol{\Omega}^2 \boldsymbol{X} \boldsymbol{Q}, \tag{5.17}$$

from which we deduce $\boldsymbol{X} \boldsymbol{Q} \boldsymbol{\Omega}^2 = \boldsymbol{\Omega}^2 \boldsymbol{X} \boldsymbol{Q}$. This relation is satisfied provided we choose $\boldsymbol{X} \boldsymbol{Q} = \boldsymbol{1}$, an identity matrix. In other words the desired transformation matrix $\boldsymbol{X}$ for the normal coordinates are nothing other than the inverse of the matrix corresponding to the normal mode solutions: $\boldsymbol{X} = \boldsymbol{Q}^{-1}$. The diagonal mass matrix is found from

$$\boldsymbol{M} = \boldsymbol{Q}^T \boldsymbol{T} \boldsymbol{Q}. \tag{5.18}$$

In summary, relation between a collection of normal coordinates $\boldsymbol{\theta}$ and a collection of ordinary coordinates $\boldsymbol{q}$ is offered by the linear transformation

$$\boldsymbol{\theta} = \boldsymbol{X}\boldsymbol{q} = \boldsymbol{Q}^{-1}\boldsymbol{q}. \tag{5.19}$$

On inverting the $\boldsymbol{Q}$ matrix we find an equivalent relation

$$\boldsymbol{q} = \boldsymbol{Q}\boldsymbol{\theta},\ \dot{\boldsymbol{q}} = \boldsymbol{Q}\dot{\boldsymbol{\theta}}. \tag{5.20}$$

When we insert this expression into the Lagrangian

$$L = \frac{1}{2}\dot{\boldsymbol{q}}^\dagger \boldsymbol{T}\dot{\boldsymbol{q}} - \frac{1}{2}\boldsymbol{q}^\dagger \boldsymbol{V}\boldsymbol{q} \tag{5.21}$$

we once again verify the normal form of the Lagrangian

$$L = \frac{1}{2}\dot{\boldsymbol{\theta}}^\dagger \boldsymbol{M}\dot{\boldsymbol{\theta}} - \frac{1}{2}\boldsymbol{\theta}^\dagger \boldsymbol{M}\boldsymbol{\Omega}^2\boldsymbol{\theta}. \tag{5.22}$$

An example of coupled harmonic oscillation is that of a linear triatomic molecule with the Lagrangian

$$L = \frac{1}{2}m\dot{x}_1^2 + \frac{1}{2}m\dot{x}_3^2 + \frac{1}{2}M\dot{x}_2^2 - V(x_1 - x_2) - V(x_2 - x_3). \tag{5.23}$$

Modeling the inter-atomic potential as a harmonic potential with the spring constant k, $V(x_1 - x_2) \approx (k/2)(x_1 - x_2)^2$, we get the equation of motion

$$\begin{pmatrix} m\ddot{x}_1 \\ M\ddot{x}_2 \\ m\ddot{x}_3 \end{pmatrix} = -k \begin{pmatrix} x_1 - x_2 \\ (x_2 - x_1) + (x_2 - x_3) \\ x_3 - x_2 \end{pmatrix}. \tag{5.24}$$

Putting it in the form $\omega^2\boldsymbol{x} = \boldsymbol{F}\boldsymbol{x}$ gives the force matrix

$$\boldsymbol{F} = \begin{pmatrix} k/m & -k/m & 0 \\ -k/M & 2k/M & -k/M \\ 0 & -k/m & k/m \end{pmatrix}. \tag{5.25}$$

There are three eigenvalues (and eigenvectors) to this:

- $\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, with the eigenvalue $\omega_1^2 = 0$. This is just an overall translation. Although this is not an oscillation mode, it still appears as one of the solutions of the normal modes. This kind of solution can be discarded in consideration of the normal vibration modes of the molecule.

- $\boldsymbol{x}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, with the eigenvalue $\omega_2^2 = k/m$. The outer two atoms oscillate out of phase, while the middle atom remains stationary. One can understand why the frequency is that of a single oscillator.

- $\boldsymbol{x}_3 = \begin{pmatrix} 1 \\ -2m/M \\ 1 \end{pmatrix}$, with the eigenvalues $\omega_3^2 = (k/m)(1+2m/M)$. The two outer atoms moves in phase, but against the motion of the inner atom. The center of mass remains stationary.

The general motion is a linear superposition of all three normal modes.

5.1.2 One-dimensional Chain with Identical Mass

A solid is a solid because the positive ions stay together despite their huge Coulomb repulsion, because the attraction with the surrounding electrons is just enough to overcome the ion-ion repulsion. The ions maintain a certain distance because that's when the total energy of the solid becomes the lowest. At T=0 the solid under consideration exists as the quantum-mechanical ground state for the Hamiltonian which defines the solid.

When the temperature is raised to a finite value, it is not only the ground state that's realized, but a lot of other low-energy excitations which lie within the range of the thermal energy T. To understand the properties of a solid, it is essential to know the whole set of allowed low-energy excitations in a solid. Among these, we want to focus on the low-energy dynamics of the ions known as the phonons.

At $T = 0$ the ions maintain a strict distance corresponding to the minimum of the total energy. That means that if the ionic separation were a little greater or a little less than this optimal value, one would see the increase of the total energy. Write the displacement vector of each ion from its optimal position by $\mathbf{u}_i$, where i indicates each ionic position in equilibrium. It is sensible to assume that the total energy depends on this displacement vector as $E(\{\mathbf{u}_i - \mathbf{u}_j\})$ for all pairs of ions at i and j. It is a function of the difference only, because moving the two ions together preserves the relative distance.

Further-neighbor ions are less tightly coupled than the near-neighbor ones, and the simplest model for the dependence of the total energy on the ionic site is achieved if we assume that it depends on the relative positions of the nearest-neighbor ions only. Assuming further that the displacements are a small fraction of the average spacing, one can expand the total energy in small displacement as

$$E \approx E_0 + \frac{K}{2} \sum_{\langle ij \rangle} (\mathbf{u}_i - \mathbf{u}_j)^2. \tag{5.26}$$

A quantum theory of the lattice dynamics would start from a Hamiltonian such as

$$H = \sum_i \frac{\mathbf{p}_i^2}{2m} + \frac{K}{2}\sum_{\langle ij\rangle}(\mathbf{u}_i - \mathbf{u}_j)^2, \tag{5.27}$$

with the commutation relations $[p_{i\alpha}, u_{j\beta}] = -i\delta_{ij}\delta_{\alpha\beta}$ ($\hbar \equiv 1$). It is possible to work out the solution of the Hamiltonian exactly, because the interaction is *harmonic*. It helps greatly to understand the classical version of the problem first because much of the classical solution carries over to the quantum case and the classical problem is conceptually simpler to treat.

As a way to build up our technical ability, it is advisable to consider the one-dimensional problem first. That is, we consider the case of a linear chain of identical mass m, each coupled to its neighboring mass to the left and right by the spring of stiffness k. The force exerted on the i-th mass is given by $k(x_{i+1} - x_i) + k(x_{i-1} - x_i) = k(x_{i+1} + x_{i-1} - 2x_i)$. Equation of motion of the i-th mass is thus

$$m\frac{d^2x_i}{dt^2} = k(x_{i+1} + x_{i-1} - 2x_i). \tag{5.28}$$

Very general consideration of the previous section suggests that a normal mode solution of the form $x_i(t) = A_i e^{-i\omega t}$ must exist, with the coefficients A_i's obeying the matrix equation

$$-m\omega^2 A_i = k(A_{i+1} + A_{i-1} - 2A_i) \quad (i = 1, \cdots, N). \tag{5.29}$$

The matrix nature of this equation become apparent when we organize the coefficients as a single, N-dimensional vector $\mathbf{A} = (A_1, \cdots, A_N)^T$. One further makes an important assumption $A_{N+1} = A_1$ and $A_0 = A_N$, equivalent to placing the N oscillators around a closed ring. Such condition is called the periodic boundary condition. Supplemented with the pbc, the above equation has an interesting property that its overall matrix form does not change under the translation of all indices $A_i \to A_{i+1}$. Such system is said to possess translation invariance. This kind of problem is known to have an exact solution given by

$$A_i = A_0 e^{ikn_i}, \tag{5.30}$$

where n_i is an integer taking on the integer value i for each site i. To check that this is the correct solution, first substitute it into the equation, and one will find

$$\begin{aligned} -m\omega^2 A_i &= k(e^{ik} + e^{-ik} - 2)A_i = -2k(1 - \cos k)A_i \\ \omega^2 &= \frac{k}{m}\left(2\sin\frac{k}{2}\right)^2. \end{aligned} \tag{5.31}$$

So we obtained that the eigenfrequency ω is equal to $2\omega_0 \cos(k/2)$. But this is not yet good, unless we know what the k value is. To learn how to determine

k, go back to the assignment $A_i = e^{ikn_i}$. We assumed that $A_{N+1} = A_1$, and in order to meet this requirement we must also assume that $e^{ikN} = 1$. A discrete set of k values are obtained from this condition, and thus a discrete set of ω_n.

$$k = \frac{2\pi n}{N}, \quad \omega_n = 2\omega_0 \sin\left(\frac{\pi n}{N}\right). \tag{5.32}$$

Counting how many different eigenfrequencies are possible, one finds there are N distinct ω_n's in the problem, the same as the number of masses on a string. And that's the right result.

The eigenfrequency spectrum we obtain through this simple exercise is in fact the exact eigenenergy spectrum of the corresponding quantum problem. The eigenenergy is the frequency ω_k multiplied by Planck's constant, $\hbar\omega_k$. With the neutron scattering one can measure the energy and momentum relation of the phonons. For realistic solids the phonon dynamics is much more complicated than the simple formula ω_k indicates. Among other reasons, this is due to the fact that there exists more than one sort of ion per unit cell in real materials. The complications brought about by multiple ions in a unit cell is illustrated by the following simple example.

5.1.3 One-dimensional Chain with Two Masses

Take the one-dimensional chain of ions again, but this time a pair of adjacent ions stay closer than between the pairs. The spring constant is similarly modified. For convenience label the left portion of the pair by an odd site $2i + 1$, and the right portion by an even site, $2i$. The dynamics of an odd site where each ion has a mass m_1 is given by

$$m_1 \frac{d^2 x_{2i+1}}{dt^2} = K_1(x_{2i+2} - x_{2i+1}) + K_2(x_{2i} - x_{2i+1}). \tag{5.33}$$

There are two spring constants K_1 and K_2 corresponding to a shorter bond and a longer one. The dynamics of an even site of mass m_2 is given by

$$m_2 \frac{d^2 x_{2i}}{dt^2} = K_2(x_{2i+1} - x_{2i}) + K_1(x_{2i-1} - x_{2i}). \tag{5.34}$$

The chain consists of N pairs of ions, and the periodic boundary condition is $x_{i+2N} = x_i$. Looking at the even sites alone, it appears that all ions are identical. The same is true of all the odd sites. Therefore, one can try out the same type of solution

$$\begin{aligned} x_{2i} &= Ae^{ik(2n_i) - i\omega t}, \\ x_{2i+1} &= Be^{ik(2n_i+1) - i\omega t}. \end{aligned} \tag{5.35}$$

In general there is no reason to expect $A = B$ because the even and odd ions are, after all, not identical. Inserting these trial solutions into the equation (5.33) and (5.34),

$$\begin{aligned} -m_2\omega^2 A &= K_2(e^{ik}B - A) + K_1(Be^{-ik} - A) \\ -m_1\omega^2 B &= K_1(e^{ik}A - B) + K_2(e^{-ik}A - B). \end{aligned} \tag{5.36}$$

After a little re-arrangement we get

$$\begin{pmatrix} K_1 + K_2 - m_2\omega^2 & -K_2e^{ik} - K_1e^{-ik} \\ -K_1e^{ik} - K_2e^{-ik} & K_1 + K_2 - m_1\omega^2 \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = 0. \tag{5.37}$$

We postpone the solution of this eigenvalue problem to an exercise. Meanwhile, we show how to solve the problem assuming $m_1 = m_2 = m$.

The characteristic equation, $(m\omega^2 - K_2 - K_1)^2 - (K_2^2 + K_1^2 + 2K_1K_2\cos 2k) = 0$ offers two solution

$$m\omega_\pm^2 = K_1 + K_2 \pm \sqrt{K_1^2 + K_2^2 + 2K_1K_2\cos 2k}. \tag{5.38}$$

The upper branch ω_+ and the lower branch ω_- are separated by a gap, which reaches a minimum value at $k = \pi/2$ as one can see from

$$\begin{aligned} m\omega_+^2 - m\omega_-^2 &= 2\sqrt{K_1^2 + K_2^2 + 2K_1K_2\cos 2k} \\ &= 2\sqrt{(K_1 - K_2)^2 + 4K_1K_2(\cos k)^2}. \end{aligned} \tag{5.39}$$

The frequency gap $\min(\omega_+ - \omega_-)$ means the existence of the frequency range in which no normal mode solutions exists.

After solving a challenging problem such as this one has to do a little sanity check. For $k = K$ one has to recover the prior solution, so let's see if this is the case. With $K_1 = K_2 = K$, Eq. (5.38) reduces to

$$m\omega^2 = 2K \pm \sqrt{2K^2(1 + \cos 2k)} = 2K(1 \pm |\cos k|). \tag{5.40}$$

In the previous case we obtain $m\omega^2 = 2K(1 - \cos k)$.

Here the separation of a single energy spectrum into two branches is artificial. But for the diatomic model the separation is real, and one can really observe two independent branches of phonons per each k known as acoustic(lower) and optical(higher) branches.

5.2 Anharmonic Oscillation

Suppose the oscillation of a particle is described by the Lagrangian

$$L = m\left(\frac{1}{2}\dot{x}^2 - \omega_0^2x^2 - \frac{1}{3}\alpha x^3 - \frac{1}{4}\beta x^4\right). \tag{5.41}$$

The resulting equation of motion is

$$\ddot{x} + \omega_0^2 x = -\alpha x^2 - \beta x^3. \tag{5.42}$$

When α and β are both very small, we can take the series solution $x = x^{(0)} + x^{(1)} + x^{(2)} + \cdots$ where the first term is $x^{(0)} = a\cos\omega t$. The frequency ω itself can be found perturbatively, $\omega = \omega_0 + \omega^{(1)} + \omega^{(2)} + \cdots$. The method, in principle, works at successive higher order and we can first try out $x = x^{(0)} + x^{(1)}$ and $\omega = \omega_0 + \omega^{(1)}$ in the equation. We will get

$$\begin{aligned} \ddot{x^{(1)}} + \omega_0^2 x^{(1)} &= -\alpha a^2(\cos\omega t)^2 + 2a\omega_0\omega^{(1)}\cos\omega t \\ &= -\frac{1}{2}\alpha a^2 - \frac{1}{2}\alpha a^2\cos 2\omega t + 2a\omega_0\omega^{(1)}\cos\omega t. \end{aligned} \tag{5.43}$$

The quartic term is dropped at this order of approximation. The r.h.s. provides the driving force for x^2, one at frequency ω and the other at 2ω. However, the one at ω is unacceptable because that was the frequency of the lowest-order solution $x^{(0)}$. To find a consistent solution we must require $\omega^{(1)} = 0$. The remaining terms can be solved to give

$$x^{(1)} = -\frac{\alpha a^2}{2\omega_0^2} + \frac{\alpha a^2}{6\omega_0^2}\cos 2\omega t = \frac{a}{2}\cdot\frac{\alpha a}{\omega_0^2}\cdot\left(-1 + \frac{1}{3}\cos 2\omega t\right). \tag{5.44}$$

By isolating the amplitude a in front, we are able to see that $\alpha a/\omega_0^2$ is a dimensionless parameter. The perturbation approach will work best when this value is small.

At next order one can put $x = x^{(0)} + x^{(1)} + x^{(2)}$ ($x^{(0)}$ and $x^{(1)}$ are already known) and $\omega = \omega_0 + \omega^{(2)}$ and find

$$\ddot{x^{(2)}} + \omega_0^2 x^{(2)} = a^3\left[\frac{1}{4}\beta - \frac{\alpha^2}{6\omega_0^2}\right]\cos 3\omega t + a\left[2\omega_0\omega^{(2)} + \frac{5a^2\alpha^2}{6\omega_0^2} - \frac{3}{4}a^2\beta\right]\cos\omega t. \tag{5.45}$$

Again the coefficient in front of $\cos\omega t$ must vanish, giving up

$$\omega^{(2)} = \omega_0\left(\frac{3\beta a^2}{8\omega_0^2} - \frac{5}{12}\left(\frac{\alpha a}{\omega_0^2}\right)^2\right). \tag{5.46}$$

The remaining driving force at the third order gives

$$x^{(2)} = \frac{a}{16}\left(\frac{1}{3}\left(\frac{\alpha a}{\omega_0^2}\right)^2 - \frac{\beta a^2}{2\omega_0^2}\right)\cos 3\omega t. \tag{5.47}$$

From the two expressions for $\omega^{(2)}$ and $x^{(2)}$ one reads off a second dimensionsless quantity, $\beta a^2/\omega_0^2$.

5.3 Motion in a Smooth Potential and Rapidly Oscillating Field

A particle moving in one dimension is subjected to a smooth potential $V(x)$ and a fast, position-independent oscillating force

$$f(x,t) = f(x)\cos\omega t. \tag{5.48}$$

By fast we mean $\omega T \gg 1$ with T the period of the motion without the fast component. For convenience we are assuming the space and time dependence can be factorized as above. Equation of motion becomes

$$m\ddot{x} = -\frac{dV}{dx} + f(x,t). \tag{5.49}$$

One can imagine that the motion will be a combination of a smooth path superimposed with a fast "jiggle". Accordingly we can represent $x(t)$ as a sum

$$x(t) = X(t) + \xi(t), \tag{5.50}$$

where $\xi(t)$ reflects the influence of the fast component. It is understood that the mean value of $\xi(t)$ over the period $2\pi/\omega$ is zero and hence $\langle x\rangle = \langle X\rangle$ over the same time period. The objective now is to describe the trajectory of the smooth component $X(t)$. Inserting $x = X + \xi$ into the equation of motion gives

$$m\ddot{X} + m\ddot{\xi} = -\frac{dV}{dX} - \xi\frac{d^2V}{dX^2} + f(X,t) + \xi\frac{\partial f(X,t)}{\partial X}. \tag{5.51}$$

We can equate the fast and slow components from both sides of the equation.

$$\begin{aligned} m\ddot{\xi} &= f(X,t) - \xi\frac{d^2V}{dX^2}, \\ m\ddot{X} &= -\frac{dV}{dX} + \xi\frac{\partial f}{\partial X}. \end{aligned} \tag{5.52}$$

In the first equation $\xi(d^2V/dX^2)$ is higher order of smallness and will be ignored. Since $f(X,t) = f(X)\cos\omega t$, we can write the solution ξ as

$$\xi = -\frac{f(X,t)}{m\omega^2}. \tag{5.53}$$

Now having found ξ, we can tackle the second equation of motion which now reads

$$m\ddot{X} = -\frac{dV}{dX} - \frac{f(X,t)}{m\omega^2}\frac{\partial f(X,t)}{\partial X} = -\frac{dV}{dX} - \frac{f(X)}{m\omega^2}\frac{\partial f(X)}{\partial X}\cos^2\omega t. \tag{5.54}$$

By integrating over one fast period we can replace $\cos^2\omega t$ by $1/2$, and then obtain

$$m\ddot{X} = -\frac{d}{dX}\left(V(X) + \frac{1}{4m\omega^2}f^2(X)\right). \tag{5.55}$$

The slow path X follows a trajectory according to the effective potential

$$V_{\text{eff}}(X) = V(X) + \frac{1}{4m\omega^2}f^2(X). \tag{5.56}$$

5.4 Problem

1. Verify that Eq. (5.9) can indeed be collected as a matrix equation (5.16).

2. Verify that Eq. (5.22) indeed follows from Eq. (5.21) upon inserting Eq. (5.20).

3. Verify that the three normal modes associated with Eq. (5.25) are the ones given in the text.

4. A pendulum consists of a mass m at the end of light rod of length l. The pivot of the pendulum is attached to a mass M which is free to slide without friction along a horizontal rail. Take the generalized coordinates to be the position x of the pivot and the angle θ that the pendulum makes with the vertical.

   a. Write down the Lagrangian and derive the equations of motion.

   b. Find the non-zero frequency of small oscillations around the stable equilibrium.

   c. Now suppose a force acts on the mass M causing it to travel with constant acceleration a in the positive-x direction. Find the equilibrium angle θ of the pendulum.

5. N equal masses are joined by $N+1$ identical springs with both ends fixed to the wall. Derive the normal mode frequencies. (Hint: You could try solving the problem with periodic boundary conditions first.)

6. Solve the two-dimensional coupled harmonic oscillator with alternating masses m_A and m_B on a square lattice. Identical spring constant K is assumed for all the neighboring masses. Write down the equation of motion, and make an appropriate ansatz for the eigenmode. Derive the eigenfrequency as a function of the two-dimensional wave vector $\mathbf{k} = (k_x, k_y)$. Plot the two dispersions and show that one of them is acoustic, the other optical.

7. Solve the one-dimensional coupled harmonic oscillator consisting of three alternating masses, m_A, m_B and $_C$. Derive the frequency spectrum and prove that only one branch is acoustic. Use the same spring constant K for all the neighboring masses. Calculate the frequency of the two optical modes at zero wave vector. Use computers if necessary to obtain your answers.

8. Derive Eq. (5.45).

9. A pendulum is described by the Lagrangian

$$L = \frac{1}{2}ml^2\dot{\theta}^2 + mgl\cos\theta. \tag{5.57}$$

The point of support of the pendulum is oscillating vertically, described by an additional force term $f(\theta, t) = -ml\omega^2 A\sin\theta\cos\omega t$. Derive the effective potential and analyze the motion. Can you find a new equilibrium position other than $\theta = 0$? Solve the problem when the support is oscillating horizontally, with an additional force $f(\theta, t) = ml\omega^2 A\cos\theta\cos\omega t$.

10. Find the characteristic equation of the matrix (5.37) and solve it. Plot the two dispersion curves you obtain.

Chapter 6

Rotation of Rigid Bodies

6.1 Kinematics

A particular point on a rigid body $\mathbf{r}(t)$ executes a rotational motion about some center position. Because this is a rigid-body rotation, it doesn't matter which point in the body we choose to describe it. The coordinates may be expressed in a certain orthogonal frame fixed to space, $\hat{e}'_\alpha$,

$$\mathbf{r}(t) = x'_\alpha(t)\hat{e}'_\alpha. \tag{6.1}$$

Alternatively, the orthogonal frame may be fixed to the rigid body itself, so that

$$\mathbf{r}(t) = x_\alpha \hat{e}_\alpha(t). \tag{6.2}$$

The coordinates x_α are independent of time. In this case, all the information about the rigid body dynamics would be contained in the equation of motion of the basis vectors $\hat{e}_\alpha(t)$. Let's see how one can derive the appropriate dynamical equations of the rigid body rotation.

Once we write

$$\hat{e}_\alpha(t) = (\hat{e}_\alpha(t) \cdot \hat{e}'_\beta)\hat{e}'_\beta = R_{\alpha\beta}(t)\hat{e}'_\beta, \tag{6.3}$$

time evolution of $\hat{e}_\alpha(t)$ follows from

$$\frac{d\hat{e}_\alpha(t)}{dt} = \frac{dR_{\alpha\beta}(t)}{dt}\hat{e}'_\beta = \frac{dR_{\alpha\beta}(t)}{dt}(R^T)_{\beta\gamma}\hat{e}_\gamma(t) = (\dot{R}R^T)_{\alpha\gamma}\hat{e}_\gamma(t). \tag{6.4}$$

This new matrix, $\mathbf{\Gamma} = \dot{\boldsymbol{R}}\boldsymbol{R}^T$, is anti-symmetric,

$$\Gamma_{\alpha\beta} = -\Gamma_{\beta\alpha}, \tag{6.5}$$

which follows from the relation $d(\boldsymbol{R}\boldsymbol{R}^T)/dt = 0$. It means the elements of the Γ matrix can be arranged as

$$\mathbf{\Gamma} = \begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix}. \tag{6.6}$$

Using the totally anti-symmetric tensor simplifies it to $\Gamma_{\alpha\beta} = \varepsilon_{\alpha\beta\gamma}\omega_\gamma$. Based on this representation of the Γ matrix we can re-write Eq. (6.4) as

$$\frac{d\hat{e}_\alpha(t)}{dt} = \varepsilon_{\alpha\beta\gamma}\omega_\gamma \hat{e}_\beta(t). \tag{6.7}$$

Expressing this relation for each α we obtain, in the case of $\alpha = 1$,

$$\frac{d\hat{e}_1(t)}{dt} = \omega_3 \hat{e}_2(t) - \omega_2 \hat{e}_3(t). \tag{6.8}$$

Cyclic permutation of the indices give the other two equations. At this point we note that the three basis vectors in the rotating frame form an orthonormal set, satisfying $\hat{e}_2(t) = \hat{e}_3(t) \times \hat{e}_1(t)$, $\hat{e}_3(t) = \hat{e}_1(t) \times \hat{e}_2(t)$. We can re-write the above equation as

$$\begin{aligned} \frac{d\hat{e}_1(t)}{dt} &= \omega_3 \hat{e}_3(t) \times \hat{e}_1(t) + \omega_2 \hat{e}_2(t) \times \hat{e}_1(t) \\ &= (\omega_1 \hat{e}_1(t) + \omega_2 \hat{e}_2(t) + \omega_3 \hat{e}_3(t)) \times \hat{e}_1(t). \end{aligned} \tag{6.9}$$

In the second line we inserted an extra factor $\omega_1 \hat{e}_1(t) \times \hat{e}_1(t) = 0$. Now it is clear that if we define the angular velocity vector $\boldsymbol{\omega}(t)$ in the rotating basis as $\boldsymbol{\omega}(t) = \omega_\alpha(t)\hat{e}_\alpha(t)$, the equation of motion takes a very simple form

$$\frac{d\hat{e}_1(t)}{dt} = \boldsymbol{\omega}(t) \times \hat{e}_1(t). \tag{6.10}$$

Obviously this sort of relation will hold for the other two indices, $\alpha = 2, 3$. Any point on the rigid body has the coordinate $\mathbf{r}(t) = x_\alpha \hat{e}_\alpha(t)$, which obeys the equation

$$\dot{\mathbf{r}} = \boldsymbol{\omega} \times \mathbf{r}. \tag{6.11}$$

This is the desired equation for the rigid body motion. Using the $\mathbf{\Gamma}$ matrix we can equally well write this as a matrix equation

$$\dot{\mathbf{r}} = \mathbf{\Gamma}(t)\mathbf{r}. \tag{6.12}$$

In general, $\boldsymbol{\omega}(t)$ (or $\mathbf{\Gamma}(t)$) depends on all sorts of external forces and changes over time in magnitude and in direction. An important property of this equation, which holds irrespective of the exact nature of $\boldsymbol{\omega}(t)$, is that the norm $\mathbf{r} \cdot \mathbf{r}$ is preserved:

$$\frac{d}{dt}(\mathbf{r}\cdot\mathbf{r}) = 2\mathbf{r}\cdot\dot{\mathbf{r}} = 2\mathbf{r}\cdot(\boldsymbol{\omega}\times\mathbf{r}) = 0. \tag{6.13}$$

Contrary to the Newtonian equation of motion for $\mathbf{r}$ which is second order in time, the one we just derived for the rigid body rotation is first-order in time, and furthermore contains the same variable $\mathbf{r}$ to linear order on both sides. This fact suggests a solution of the form

$$\mathbf{r}(t) = \exp\Big(\int_0^t \boldsymbol{\Gamma}(t')dt'\Big)\mathbf{r}(0). \tag{6.14}$$

If $\boldsymbol{\Gamma}(t)$ were a constant $\boldsymbol{\Gamma}(t) = \boldsymbol{\Gamma}$ the integral simplifies to

$$\exp\Big(\int_0^t \boldsymbol{\Gamma}(t')dt'\Big) = \exp\Big(t\boldsymbol{\Gamma}\Big) = \sum_{n=0}^{\infty}\frac{1}{n!}(t\boldsymbol{\Gamma})^n. \tag{6.15}$$

The meaning of the exponential operation becomes much more tricky when $\boldsymbol{\Gamma}(t)$ is time-dependent. In fact the correct way to write the solution to Eq. (6.12) is

$$\mathbf{r}(t) = \sum_{n=0}^{\infty}\int_0^t dt_1\int_0^{t_1} dt_2\cdots\int_0^{t_{n-1}} dt_n\ \boldsymbol{\Gamma}(t_1)\boldsymbol{\Gamma}(t_2)\cdots\boldsymbol{\Gamma}(t_n)\mathbf{r}(0). \tag{6.16}$$

We are choosing $0 < t_n < t_{n-1} < \cdots < t_2 < t_1 < t$ by appropriately limiting the upper bound of the integration for each time piece t_i. The Γ matrices are arranged such that those at later times always appear to the left of those acting at earlier times. It can be verified easily by differentiation that the above expression satisfies $\dot{\mathbf{r}} = \boldsymbol{\Gamma}(t)\mathbf{r}$.

We must emphasize that the Schr$\acute{}$odinger equation itself can be cast in the similar form (take $\hbar \equiv 1$)

$$\dot{\psi} = -iH(t)\psi. \tag{6.17}$$

For an arbitrary time-dependent Hamiltonian $H(t)$, the wave function at a later time t can be found from the initial wave function $\psi(0)$ by the same equation (6.16), only by replacing the matrix $\omega(t_i)$ by the Hamiltonian $-iH(t_i)$. We see that much of the mathematical structure of quantum mechanics is already manifest in a classical problem of spinning top!

A second strong hint of the direct analogy between the classical spinning top and the quantum mechanics follows from analyzing the kinetic energy expression for a top:

$$T = \frac{1}{2}\sum_i m_i\dot{\mathbf{r}}_i{}^2 = \frac{1}{2}\sum_i m_i(\boldsymbol{\omega}\times\mathbf{r}_i)^2. \tag{6.18}$$

In a rigid body every point $\mathbf{r}_i$ is subject to the same equation of motion derived earlier, $\dot{\mathbf{r}}_i = \boldsymbol{\omega}\times\mathbf{r}_i$. If we express both vectors $\boldsymbol{\omega}$ and $\mathbf{r}_i$ in the body-fixed frame,

$$\begin{aligned} \boldsymbol{\omega}(t) &= \omega_\alpha(t)\hat{e}_\alpha(t), \\ \mathbf{r}_i(t) &= x_{i,\alpha}\hat{e}_\alpha(t), \end{aligned} \tag{6.19}$$

the cross product becomes

$$\boldsymbol{\omega}(t) \times \mathbf{r}_i(t) = \omega_\alpha(t) x_{i,\beta}\hat{e}_\alpha(t) \times \hat{e}_\beta(t) = \varepsilon_{\alpha\beta\gamma}\omega_\alpha(t)x_{i,\beta}\hat{e}_\gamma(t). \tag{6.20}$$

Squaring this and summing over all the particles,

$$\begin{aligned} \sum_i (\boldsymbol{\omega}(t) \times \mathbf{r}_i(t))^2 &= \sum_i \varepsilon_{\alpha\beta\gamma}\varepsilon_{\alpha'\beta'\gamma'}\omega_\alpha(t)\omega_{\alpha'}(t)x_{i,\beta}x_{i,\beta'}\big(\hat{e}_\gamma(t)\cdot\hat{e}_{\gamma'}(t)\big) \\ &= \sum_i \big(\delta_{\alpha\alpha'}\delta_{\beta\beta'} - \delta_{\alpha\beta'}\delta_{\beta\alpha'}\big)\omega_\alpha(t)\omega_{\alpha'}(t)x_{i,\beta}x_{i,\beta'} \\ &= \omega_\alpha(t)\Big(\delta_{\alpha\beta}\sum_i x_{i,\alpha}x_{i,\beta} - \sum_i x_{i,\alpha}x_{i,\beta}\Big)\omega_\beta(t). \end{aligned} \tag{6.21}$$

The expression inside the parenthesis is independent of time, and reflects the intrinsic nature of the rigid body itself. The inertia tensor $I_{\alpha\beta}$, with α and β running over the three axes $1, 2, 3$ are given by

$$I_{\alpha\beta} = \delta_{\alpha\beta}\sum_i x_{i,\alpha}x_{i,\beta} - \sum_i x_{i,\alpha}x_{i,\beta} = \int d^3\mathbf{x}\ \rho(\mathbf{x})\Big(\mathbf{x}\cdot\mathbf{x}\delta_{ab} - (\mathbf{x})_a(\mathbf{x})_b\Big). \tag{6.22}$$

We are using the vector notation $\mathbf{x} = (x_1, x_2, x_3)$ to denote the three components of the position vector $\mathbf{r}(t) = x_\alpha\hat{e}_\alpha(t)$ in the body-fixed frame. With the help of the inertial tensor one can write the kinetic energy of a rotating rigid body as

$$T(t) = \frac{1}{2}\mathbf{w}^T(t)I\mathbf{w}(t), \quad \mathbf{w}(t) = \begin{pmatrix}\omega_1(t)\\ \omega_2(t)\\ \omega_3(t)\end{pmatrix}. \tag{6.23}$$

We are using a different notation $\mathbf{w}(t)$ for a collection of components of the angular velocity vector $\boldsymbol{\omega}(t)$. It is important to keep in mind that the collection of coefficients $\mathbf{w}(t)$ is not the same as the true angular velocity vector $\boldsymbol{\omega}(t)$.

The inertia tensor I_{ab} is symmetric, $I_{ab} = I_{ba}$, and real, and is a subset of the class of matrices known as Hermitian matrices. As such, they always possess three sets of orthogonal eigenvectors $\hat{E}_\alpha$ and real eigenvalues I_α which obey

$$I\hat{E}_\alpha = I_\alpha\hat{E}_\alpha, \quad \hat{E}_\alpha\cdot\hat{E}_\beta = \delta_{\alpha\beta}. \tag{6.24}$$

The three axes derived this way are called the principal axes of rotation.

Let's check out the meaning of the three principal axis vectors just derived. We know that the inertia tensor matrix I can be reconstructed in terms of its eigenvectors and eigenvalues as

$$I = \sum_\alpha I_\alpha \hat{E}_\alpha \hat{E}_\alpha^T. \tag{6.25}$$

That is, by writing each $\hat{E}_\alpha$ as a column vector and its transpose $\hat{E}_\alpha^T$ as a row vector, one obtain a matrix $\hat{E}_\alpha \hat{E}_\alpha^T$. Summing over the product of this matrix with the eigenvalue I_α over all indices α reproduces the original inertial matrix I.

With the decomposition of the inertia matrix, the kinetic energy we found in Eq. (6.23) becomes

$$T(t) = \frac{1}{2} \sum_\alpha I_\alpha \Big(\mathbf{w}^T(t) \hat{E}_\alpha \Big) \Big(\hat{E}_\alpha^T \mathbf{w}(t) \Big). \tag{6.26}$$

Now, each term inside the parenthesis $\hat{E}_\alpha^T \mathbf{w}(t)$ is a scalar. In fact this is the inner product of the instantaneous angular velocity vector $\boldsymbol{\omega}(t)$ with the particular principal axis vector $\hat{E}_\alpha(t)$. Had we chosen the body-fixed frame to coincide with the principal axes, $\hat{e}_\alpha(t) \equiv \hat{E}(t)$, the product $\hat{E}_\alpha \mathbf{w}(t)$ would be exactly the inner product $\boldsymbol{\omega}(t) \cdot \hat{E}_\alpha(t) = \omega_\alpha(t)$, which is the component of the angular velocity vector along the α-th principal axis. With this understanding in mind, the kinetic energy of rotation becomes

$$T(t) = \frac{1}{2} \sum_\alpha I_\alpha (\omega_\alpha(t))^2. \tag{6.27}$$

If we expand the instantaneous angular velocity vector $\boldsymbol{\omega}$ not in an arbitrary body-fixed frame, but in a particular frame spanned by $\hat{E}_\alpha$, we discover that the kinetic energy is cast in a simple form given above, where the off-diagonal elements of the inertia tensor vanish. This tells us there is a natural choice of body-fixed frame of reference with which to describe the rotation of a rigid body, and in such a frame the elements of the inertia tensor only has diagonal elements.

Almost exactly the same story carries over to quantum mechanics. A particular quantum mechanical state $|\psi\rangle$ expanded in a certain basis set $|\alpha\rangle$ as

$$|\psi\rangle = \sum_\alpha |\alpha\rangle\langle\alpha|\psi\rangle = \sum_\alpha \psi_\alpha |\alpha\rangle \tag{6.28}$$

has the expectation value of the Hamiltonian

$$\langle\psi|H|\psi\rangle = \sum_{\alpha,\beta} \psi_\alpha^* H_{\alpha\beta} \psi_\beta = \boldsymbol{\psi}^\dagger \boldsymbol{H} \boldsymbol{\psi}. \tag{6.29}$$

In an arbitrarily chosen orthogonal basis $|\alpha\rangle$ the matrix elements of the Hermitian matrix $H_{\alpha\beta}$ carry non-diagonal as well as diagonal elements. $\boldsymbol{H}$, being Hermitian, always allows a set of orthogonal eigenvectors and real eigenvalues

$$\boldsymbol{H} \boldsymbol{\phi}_\alpha = \varepsilon_\alpha \boldsymbol{\phi}_\alpha, \tag{6.30}$$

allowing the re-construction of the Hamiltonian matrix as

$$\boldsymbol{H} = \sum_{\alpha} \varepsilon_{\alpha} \boldsymbol{\phi}_{\alpha} \boldsymbol{\phi}_{\alpha}^{\dagger}. \tag{6.31}$$

Expectation value can be cast in the diagonal form

$$\langle \psi | H \psi \rangle = \sum_{\alpha} \varepsilon_{\alpha} (\boldsymbol{\psi}^{\dagger} \boldsymbol{\phi}_{\alpha})(\boldsymbol{\phi}_{\alpha}^{\dagger} \boldsymbol{\psi}) = \sum_{\alpha} \varepsilon_{\alpha} |\boldsymbol{\psi}^{\dagger} \boldsymbol{\phi}_{\alpha}|^2. \tag{6.32}$$

The angular momentum vector $\mathbf{L}$ of a rotating rigid body is the sum of the angular momenta of the individual constituent particles,

$$\mathbf{L}(t) \quad = \quad \sum_{i} m_i \mathbf{r}_i(t) \times \dot{\mathbf{r}}_i(t) = \sum_{i} m_i \mathbf{r}_i(t) \times (\boldsymbol{\omega}(t) \times \mathbf{r}_i(t)). \tag{6.33}$$

We go through the same exercise as before in writing out the vectors in the body-fixed frame,

$$\begin{aligned} \boldsymbol{\omega}(t) &= \omega_{\alpha}(t) \hat{e}_{\alpha}(t), \\ \mathbf{r}_i(t) &= x_{i,\alpha} \hat{e}_{\alpha}(t), \\ \boldsymbol{\omega}(t) \times \mathbf{r}_i(t) &= \varepsilon_{\alpha\beta\gamma} \omega_{\alpha}(t) x_{i,\beta} \hat{e}_{\gamma}(t). \end{aligned} \tag{6.34}$$

It follows,

$$\begin{aligned} \mathbf{r}_i(t) \times (\boldsymbol{\omega}(t) \times \mathbf{r}_i(t)) &= \varepsilon_{\alpha\beta\gamma} \omega_{\alpha}(t) x_{i,\beta} x_{i,\delta} \hat{e}_{\delta}(t) \times \hat{e}_{\gamma}(t) \\ &= \varepsilon_{\gamma\alpha\beta} \varepsilon_{\gamma\lambda\delta} \omega_{\alpha}(t) x_{i,\beta} x_{i,\delta} \hat{e}_{\lambda}(t) \\ &= \Big(\delta_{\alpha\lambda} \delta_{\beta\delta} - \delta_{\alpha\delta} \delta_{\beta\lambda} \Big) \omega_{\alpha}(t) x_{i,\beta} x_{i,\delta} \hat{e}_{\lambda}(t) \\ &= x_{i,\beta} x_{i,\beta} \omega_{\alpha}(t) \hat{e}_{\alpha}(t) - x_{i,\alpha} x_{i,\beta} \omega_{\beta}(t) \hat{e}_{\alpha}(t). \end{aligned} \tag{6.35}$$

Component of the angular momentum vector $\mathbf{L}(t)$ along a particular body-fixed axis $\hat{e}_{\alpha}(t)$ is therefore

$$\begin{aligned} \mathbf{L}(t) \cdot \hat{e}_{\alpha}(t) = L_{\alpha}(t) &= \Big(\delta_{\alpha\beta} \sum_{i} x_{i,\beta} x_{i,\beta} - \sum_{i} x_{i,\alpha} x_{i,\beta} \Big) \omega_{\beta}(t) \\ &= I_{\alpha\beta} \omega_{\beta}(t). \end{aligned} \tag{6.36}$$

It shows that components of the angular velocity $\boldsymbol{\omega}(t)$ is not in general proportional to the components of the angular momentum vector $\mathbf{L}(t)$. The same is actually true of motion of particles in an anisotropic medium where the mass becomes a tensor quantity m_{ab}. Only the sphere or the cube will have the isotropic moment of inertia tensor $I_{\alpha\beta} = I \delta_{\alpha\beta}$.

Collecting components of the angular momentum vector $L_{\alpha}(t)$ as a column vector gives the linear matrix equation

$$\boldsymbol{L}(t) = \begin{pmatrix} L_1(t) \\ L_2(t) \\ L_3(t) \end{pmatrix} = I\mathbf{w}(t) = \sum_\alpha I_\alpha \hat{E}_\alpha \Big(\hat{E}_\alpha^T \mathbf{w}(t) \Big). \tag{6.37}$$

Previously we argued that $\hat{E}_\alpha^T \mathbf{w}(t)$ is the α-th component of the angular velocity vector in the principal axis frame. Adopting $\omega_\alpha(t) = \hat{E}_\alpha^T \mathbf{w}(t)$, the above equation becomes

$$\boldsymbol{L}(t) = \sum_\alpha I_\alpha \omega_\alpha(t) \hat{E}_\alpha. \tag{6.38}$$

We can think of this result as saying that the angular momentum vector has a component along a given principal axis $\hat{E}_\alpha$ equal to $I_\alpha \omega_\alpha(t)$. The angular velocity and the angular momentum vectors have the representation

$$\begin{aligned} \boldsymbol{\omega}(t) &= \sum_\alpha \omega_\alpha(t) \hat{E}_\alpha(t), \\ \mathbf{L}(t) &= \sum_\alpha I_\alpha \omega_\alpha(t) \hat{E}_\alpha(t). \end{aligned} \tag{6.39}$$

6.2 Euler's equation

Think of a satellite or an asteroid that travels through space while executing some sort of rotation on its own. One can express the position of the particles on such object as the sum

$$\mathbf{R}_i(t) = \mathbf{R}(t) + \mathbf{r}_i(t) \tag{6.40}$$

using the center of mass position $\mathbf{R}$, and $\dot{\mathbf{r}}_i = \boldsymbol{\omega} \times \mathbf{r}_i$. The kinetic energy of the object can be shown to separate,

$$T = \frac{1}{2} \sum_i m_i \dot{\mathbf{R}}_i^2 = \frac{1}{2} M \dot{\mathbf{R}}^2 + \frac{1}{2} \mathbf{w}^T(t) I \mathbf{w}(t). \tag{6.41}$$

For an observer traveling at the velocity, $\dot{\mathbf{R}}$, the motion of the object is entirely that of the rotation. All the tools we've developed for analysis of rotational motion apply.

For a rigid body rotating in free space (external force = 0, external torque = 0) the total angular momentum must be a conserved quantity. We can express this statement using the principal axes, $\mathbf{L} = \sum_\alpha I_\alpha \omega_\alpha(t) \hat{E}_\alpha(t)$.

$$\begin{aligned} \dot{\mathbf{L}} = 0 &= \sum_\alpha I_\alpha \dot{\omega}_\alpha \hat{E}_\alpha + \sum_\alpha I_\alpha \omega_\alpha \dot{\hat{E}}_\alpha \\ &= \sum_\alpha I_\alpha \dot{\omega}_\alpha \hat{E}_\alpha + \sum_\alpha I_\alpha \omega_\alpha \boldsymbol{\omega} \times \hat{E}_\alpha. \end{aligned} \tag{6.42}$$

Each component must vanish,

$$I_\alpha \dot{\omega}_\alpha + \sum_\beta I_\beta \omega_\beta \boldsymbol{\omega} \cdot (\hat{E}_\alpha \times \hat{E}_\beta) = I_\alpha \dot{\omega}_\alpha + \sum_\beta \varepsilon_{\alpha\beta\gamma} I_\beta \omega_\beta \omega_\gamma = 0. \tag{6.43}$$

The three relations

$$\begin{aligned} I_1 \dot{\omega}_1 &= \omega_2 \omega_3 (I_2 - I_3) \\ I_2 \dot{\omega}_2 &= \omega_3 \omega_1 (I_3 - I_1) \\ I_3 \dot{\omega}_3 &= \omega_1 \omega_2 (I_1 - I_2) \end{aligned} \tag{6.44}$$

are known as Euler's equations of a free rigid body.

If the rigid body had a perfectly symmetric inertia tensor $I_1 = I_2 = I_3 = I$ the equations reduce to $\dot{\omega}_\alpha = 0$, or $\omega_\alpha(t) = \omega_\alpha$. Angular momentum vector in this case becomes $\mathbf{L}(t) = \sum_\alpha I_\alpha \omega_\alpha \hat{E}_\alpha(t) = I \sum_\alpha \omega_\alpha \hat{E}_\alpha(t)$, so there is still some time dependence through the body-fixed basis vector $\hat{E}_\alpha(t)$. However, taking the time derivative of it and using $\dot{\hat{E}}_1(t) = \omega_2 \hat{E}_3(t) - \omega_3 \hat{E}_2(t)$, etc., gives

$$\dot{\mathbf{L}}(t) = \sum_{\alpha,\beta,\gamma} I_\alpha \omega_\alpha \varepsilon_{\alpha\beta\gamma} \omega_\beta \hat{E}_\gamma(t) = I \sum_\gamma \left(\sum_{\alpha,\beta} \varepsilon_{\alpha\beta\gamma} \omega_\alpha \omega_\beta \right) \hat{E}_\gamma(t) = 0. \tag{6.45}$$

The angular momentum vector $\mathbf{L}$ is a constant of motion, as expected.

Next comes the less trivial and more interesting case $I_1 = I_2 \neq I \neq I_3$. Euler's equations reduce to

$$\begin{aligned} I \dot{\omega}_1 &= \delta I \omega_2 \omega_3 \\ I \dot{\omega}_2 &= \delta I \omega_3 \omega_1 \\ I_3 \dot{\omega}_3 &= 0, \end{aligned} \tag{6.46}$$

the difference $\delta I = I - I_3$ being either positive or negative. Fortunately the third equation can be solved right away to give $\omega_3(t) = B$, which is fed back to the first two to give

$$\begin{aligned} \dot{\omega}_1(t) - \Omega \omega_2(t) &= 0 \\ \dot{\omega}_2(t) + \Omega \omega_1(t) &= 0 \end{aligned} \tag{6.47}$$

A new frequency scale $\Omega = (\delta I / I) B$ appears in the governing dynamics due to the asymmetry of the top. Altogether, components of the angular velocity vector is

$$(\omega_1(t), \omega_2(t), \omega_3(t)) = (A \sin \Omega t, A \cos \Omega t, B). \tag{6.48}$$

To an observer fixed to and rotating with the rigid body, the angular velocity vector $\boldsymbol{\omega}(t)$ appears to execute a precess around the body-fixed $\hat{E}_3(t)$ axis with the angular velocity $|\Omega|$. The angular momentum vector comes out to be

$$\begin{aligned}\mathbf{L} &= \sum_\alpha I_\alpha \omega_\alpha(t)\hat{E}_\alpha(t) \\ &= IA\sin(\Omega t)\hat{E}_1(t) + IA\cos(\Omega t)\hat{E}_2(t) + (I+\delta I)B\hat{E}_3(t). \qquad (6.49)\end{aligned}$$

Since $\mathbf{L}$ is a constant, the body-fixed vectors $\hat{E}_\alpha(t)$ must assume such time dependence as to cancel out the explicit time dependences in $\omega_\alpha(t)$. For example, $\hat{E}_3(t) = \hat{z}$ and

$$\begin{aligned}\hat{E}_1(t) &= \hat{x}\cos(\Omega t) + \hat{y}\sin(\Omega t), \\ \hat{E}_2(t) &= \hat{y}\cos(\Omega t) - \hat{x}\sin(\Omega t), \qquad (6.50)\end{aligned}$$

can give a constant $\mathbf{L}$. This is just the rotation of the body about the $\hat{z}$ axis.

There is a more exotic solution satisfying a constant $\mathbf{L}$ called the "wobble". I have tried to derive the wobble solution for several hours and failed. There is a neat way to show this, but it requires a new language called the Euler angle. We are going to introduce the technology of the Euler angle first, and derive the wobble solution from it.

An arbitrary pair of orthogonal frames $\{\hat{e}_\alpha\}$ and $\{\hat{e}'_\alpha\}$ are related by an orthogonal matrix such that each $\hat{e}'_\alpha$ is related to the linear combination of $\hat{e}_\beta$'s as $\hat{e}'_\alpha = R_{\alpha\beta}\hat{e}_\beta$. There are nine elements in R, but due to the orthogonality condition $R_{\alpha\beta}R_{\alpha\gamma} = \delta_{\beta\gamma}$ (there are six independent equations in this) there are really only three independent elements in any given R. The great mathematician Leonhard Euler (1707-1783) proved that any given R can be parameterized in terms of three angles ψ, θ, ϕ, as

$$\boldsymbol{R}[\psi,\theta,\phi] = \boldsymbol{R}_3[\psi]\boldsymbol{R}_1[\theta]\boldsymbol{R}_3[\phi]. \qquad (6.51)$$

The subscripts in each R on the right side indicates the axis about which the rotation is taking place. So we have

$$\begin{aligned}\boldsymbol{R}_3[\phi] &= \begin{pmatrix}\cos\phi & \sin\phi & 0\\ -\sin\phi & \cos\phi & 0\\ 0 & 0 & 1\end{pmatrix}, \\ \boldsymbol{R}_1[\theta] &= \begin{pmatrix}1 & 0 & 0\\ 0 & \cos\theta & \sin\theta\\ 0 & -\sin\theta & \cos\theta\end{pmatrix}, \\ \boldsymbol{R}_3[\psi] &= \begin{pmatrix}\cos\psi & \sin\psi & 0\\ -\sin\psi & \cos\psi & 0\\ 0 & 0 & 1\end{pmatrix}, \qquad (6.52)\end{aligned}$$

The product of three matrices have the complicated form

$$\boldsymbol{R} = \begin{pmatrix}\cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & \sin\phi\cos\psi + \cos\theta\sin\psi\cos\phi & \sin\theta\sin\psi\\ -\cos\phi\sin\psi - \cos\theta\cos\psi\sin\phi & -\sin\psi\sin\phi + \cos\theta\cos\psi\cos\phi & \sin\theta\cos\psi\\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta\end{pmatrix} \qquad (6.53)$$

We can think of $\boldsymbol{R}$ as the rotation matrix implementing the rotation of the space-fixed coordinate frame to the body-fixed frame. Since the body-fixed frame changes with time, so is the rotation matrix $\boldsymbol{R}(t)$ through time dependence of the three angles.

At the beginning of this chapter we showed that $\boldsymbol{\Gamma} = \dot{\boldsymbol{R}}\boldsymbol{R}^T$ is an anti-symmetric matrix whose components are the angular velocity $\boldsymbol{\omega}$ through the relation

$$\boldsymbol{\Gamma} = \begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix}. \tag{6.54}$$

Since we have the rotation $\boldsymbol{R}$ in terms of three Euler angles we may be able to find $\boldsymbol{\omega}$ also in terms of the Euler angles. This is tedious to do by hand, but Mathematica can find the answer in no time:

$$\begin{aligned} \omega_1(t) &= \dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi, \\ \omega_2(t) &= \dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi, \\ \omega_3(t) &= \dot{\psi} + \dot{\phi}\cos\theta. \end{aligned} \tag{6.55}$$

These are the components of the angular velocity vector along the body-fixed frame.

We go back to the problem of a free spinning top with $I_1 = I_2 = I \neq I_3$. It was already shown that $\dot{\omega}_3 = 0$ in this case, translating into

$$\dot{\psi} + \dot{\phi}\cos\theta = B \tag{6.56}$$

in terms of Euler angles. Kinetic energy is a constant of motion for the free spinning top,

$$T = \frac{1}{2}\boldsymbol{\omega}(t)\cdot\mathbf{L} = \frac{1}{2}I([\omega_1(t)]^2 + [\omega_2(t)]^2) + \frac{1}{2}I_3\omega_3^2. \tag{6.57}$$

With the constancy of ω_3 we find $[\omega_1(t)]^2 + [\omega_2(t)]^2 = A^2$ is a constant too, in agreement with the previous analysis. In terms of Euler angles, the statement is

$$(\dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi)^2 + (\dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi)^2 = \dot{\theta}^2 + \dot{\phi}^2(\sin\theta)^2 = A^2. \tag{6.58}$$

Constancy of $\boldsymbol{\omega}(t)\cdot\mathbf{L}$ means that $\boldsymbol{\omega}(t)$, while not itself a constant, executes a precession around the constant $\mathbf{L}$ vector as seen by the space-fixed observer. One can derive an interesting fact that $\mathbf{L}$, $\boldsymbol{\omega}(t)$, and $\hat{E}_3(t)$ all lie in the plane by showing

$$\mathbf{L}\cdot(\boldsymbol{\omega}(t)\times\hat{E}_3(t)) = 0. \tag{6.59}$$

First there is $\boldsymbol{\omega}(t)\times\hat{E}_3(t) = -\omega_1(t)\hat{E}_2(t) + \omega_2(t)\hat{E}_1(t)$. Then the inner product of this with $\mathbf{L} = \sum_\alpha I_\alpha(\omega_\alpha(t)\hat{E}_\alpha(t)$ gives $I\omega_2(t)(-\omega_1(t)) + I\omega_1(t)\omega_2(t) = 0$. To

an observer in the space-fixed frame, $\boldsymbol{\omega}(t)$ traces out a cone around the direction of the $\mathbf{L}$ vector. It also traces out a cone around the body-fixed $\hat{E}_3(t)$ axis. You can imagine the situation of two cylinders rolling against each other, with the orientation of one of the cones fixed in space, while the other cone rolls around the first one. The angular velocity vector $\boldsymbol{\omega}(t)$ points along the interface of the two cones, which explains why $\mathbf{L}, \boldsymbol{\omega}(t), \hat{E}_3(t)$ all lie in a plane. Further implication of Eq. (6.59) is that

$$\frac{d}{dt}\left(\mathbf{L}\cdot\hat{E}_3(t)\right) = \mathbf{L}\cdot\left(\boldsymbol{\omega}(t)\times\hat{E}_3(t)\right) = 0, \tag{6.60}$$

leading to the constancy of $\mathbf{L}\cdot\hat{E}_3(t)$. The body-fixed axis $\hat{E}_3(t)$ always makes a constant angle with the angular momentum vector $\mathbf{L}$! If we choose the constant direction of $\mathbf{L}$ to represent the space-fixed axis $\hat{e}'_3$, Euler rotation angle ϕ becomes the rotation of the frame with respect to the $\mathbf{L}$-axis and so on. In particular the angle between $\mathbf{L}$ and $\hat{E}_3(t)$ is none other than θ, which by the argument just presented must be a constant of time.

With the help of the extra relation $\theta = \theta_0$, we find the expressions for $\omega_1(t)$ and $\omega_2(t)$ simplify greatly to

$$\begin{aligned} \omega_1(t) &= \dot{\phi}\sin\theta_0\sin\psi, \\ \omega_2(t) &= \dot{\phi}\sin\theta_0\cos\psi. \end{aligned} \tag{6.61}$$

Furthermore, $[\omega_1(t)]^2 + [\omega_2(t)]^2 = \dot{\theta}^2 + \dot{\phi}^2(\sin\theta)^2 = A^2$ reduces to $\dot{\phi}^2(\sin\theta_0)^2 = A^2$, which gives the constant rate of change

$$|\dot{\phi}| = \left|\frac{A}{\sin\theta_0}\right|. \tag{6.62}$$

On substituting Eq. (6.61) into

$$\begin{aligned} \dot{\omega}_1(t) - \Omega\omega_2(t) &= 0 \\ \dot{\omega}_2(t) + \Omega\omega_1(t) &= 0, \end{aligned} \tag{6.63}$$

we find a single relation emerges: $\dot{\psi} = \Omega$. The rate of change of the third Euler angle is nothing but the precessional frequency $\dot{\psi} = \Omega = (\delta I/I)\omega_3$. Substituting $\omega_3 = \dot{\psi} + \dot{\phi}\cos\theta_0$ in this formula gives a new expression for $\dot{\phi}$:

$$\dot{\psi} = \frac{\delta I}{I}(\dot{\psi} + \dot{\phi}\cos\theta_0) \quad\rightarrow\quad \dot{\phi} = \left(\frac{I-\delta I}{\delta I}\right)\frac{\dot{\psi}}{\cos\theta_0} = \left(\frac{I_3}{I-I_3}\right)\frac{\Omega}{\cos\theta_0}. \tag{6.64}$$

If you take a rigid body in the shape of a plate, the body-fixed axis $\hat{E}_3(t)$ is roughly orthogonal to the plate. The component of $\boldsymbol{\omega}(t)$ of the plate about $\hat{E}_3(t)$ represents how fast the plate spins around about its own axis. At the same time the whole plate wobbles about the $\mathbf{L}$ axis at the rate given by $\dot{\phi}$. The ratio of the spin to the wobble, both of which are constants, is given by

$$\frac{\omega_3}{\dot{\phi}} = \frac{\Omega}{\dot{\phi}} + \cos\theta_0 = \left(\frac{I - I_3}{I_3} + 1\right)\cos\theta_0 = \frac{I}{I_3}\cos\theta_0 \approx \frac{1}{2}\cos\theta_0. \qquad (6.65)$$

For a plate we have $I_3 \approx 2I$. If the plate is spinning almost vertically, $\theta_0 \approx 0$, we can approximate $\cos\theta_0 \approx 1$ and obtain $\omega_3 \approx \dot{\phi}/2$. In other words the plate wobbles ($\dot{\phi}$) twice as fast as it spins (ω_3). This was the famous observation recorded in Feynman's book *Surely you're joking Mr. Feynman*:

I was in the cafeteria and some guy, fooling around, throws a plate in the air. As the plate went up in the air I saw it wobble, and I noticed the red medallion of Cornell on the plate going around. It was pretty obvious to me that the medallion went around faster than the wobbling.

The medallion going around is the spin. He further says

I had nothing to do, so I start figuring out the motion of the rotating plates. I discover that when the angle is very slight, the medallion rotates twice as fast as the wobble rate - two to one. It came out of a complicated equation!

The slight angle he mentions refers to $\theta_0 \approx 0$. His conclusion about the rotation of the medallion (spin) being twice as fast as the wobble rate, however, is contrary to what we have just derived! Perhaps his equations were worked out on a piece of napkin, which led to some error. Anyway, he did other things right and won a Nobel prize.

We have dealt with free spinning top so far. When a top is pinned at a certain point and acted on by gravity, the Lagrangian becomes

$$\begin{aligned} L &= \frac{1}{2}I^2(\omega_1^2 + \omega_2^2) + \frac{1}{2}I_3\omega_3^2 - Mgl\cos\theta \\ &= \frac{1}{2}I_1(\dot{\theta}^2 + \dot{\phi}^2\sin^2\theta) + \frac{1}{2}I_3(\dot{\psi} + \dot{\phi}\cos\theta)^2 - Mgl\cos\theta. \end{aligned} \qquad (6.66)$$

See how the kinetic energy is readily obtained in terms of the Euler angles!

There are two constants of motion as one can see from the absence of ψ and ϕ (only their derivatives appear) in the Lagrangian. The two constants of motion are

$$\begin{aligned} \frac{\partial L}{\partial \dot{\psi}} &= I_3(\dot{\psi} + \dot{\phi}\cos\theta) = P_\psi, \\ \frac{\partial L}{\partial \dot{\phi}} &= I\dot{\phi}\sin^2\theta + I_3(\dot{\psi} + \dot{\phi}\cos\theta)\cos\theta \\ &= I\dot{\phi}\sin^2\theta + P_\psi\cos\theta = P_\phi. \end{aligned} \qquad (6.67)$$

A little re-arrangement gives

$$\begin{aligned}\dot{\phi} &= \frac{P_\phi - P_\psi \cos\theta}{I \sin^2\theta}. \\ \dot{\psi} &= \frac{P_\psi}{I_3} - \dot{\phi}\cos\theta = \frac{P_\psi}{I_3} + \frac{P_\psi \cos\theta - P_\phi}{I}\frac{\cos\theta}{\sin^2\theta}. \end{aligned} \tag{6.68}$$

Once we know how $\theta(t)$ behave over time, the other two variables can be solved by simple integration of the above equations. So how do we solve for θ? First work out

$$\begin{aligned}\frac{\delta L}{\delta \dot{\theta}} &= I\dot{\theta}, \\ \frac{\delta L}{\delta \theta} &= I\dot{\phi}^2 \sin\theta\cos\theta - I_3\dot{\phi}(\dot{\psi} + \dot{\phi}\cos\theta)\sin\theta + Mgl\sin\theta \\ &= \frac{(P_\phi - P_\psi\cos\theta)^2}{I\sin^3\theta}\cos\theta - P_\psi\frac{P_\phi - P_\psi\cos\theta}{I\sin\theta} + Mgl\sin\theta \\ &= -\frac{d}{d\theta}\left(\frac{(P_\phi - P_\psi\cos\theta)^2}{2I\sin^2\theta} + Mgl\cos\theta\right) = -\frac{dV_{\rm eff}(\theta)}{d\theta}. \end{aligned} \tag{6.69}$$

Euler-Lagrange equation is $I\ddot{\theta} + V'_{\rm eff}(\theta) = 0$. The motion is effectively one-dimensional, with the range of θ confined to the region $\theta_1 < \theta < \theta_2$ bounded by two end points $V_{\rm eff}(\theta_1) = E = V_{\rm eff}(\theta_2)$.

6.3 Problem

1. Prove Eq. (6.5).
2. Show that Eq. (6.16) satisfies $\dot{\mathbf{r}} = \mathbf{\Gamma}(t)\mathbf{r}$.
3. Prove Eq. (6.41).
4. Derive Eq. (6.53) by hand.
5. Derive Eq. (6.55) from Eq. (6.53) using Mathematica. Please include a copy of your Mathematica program.
6. Show that for any solid, the sum of any two principal moments of inertia is not less than the third. For what shapes is the sum of two equal to the third?
7. Calculate the moments of inertia of

   a. A uniform sphere of mass M, radius R

   b. A hollow sphere of mass M, radius R

   c. A uniform circular cone of mass M, height H, and base radius R, with respect to the principal axes whose origin is at the vertex of the cone

   d. A solid uniform cylinder of radius R, height $2H$, and mass M, with respect to its centre of mass. For what height-to-radius ratio does the cylinder spin like a sphere?

e. A uniform ellipsoid of mass M, defined by

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq R^2, \tag{6.70}$$

with respect to the (x, y, z) axes with origin at the centre of mass. (Hint: With a change of coordinates you can reduce this problem to that of the solid sphere.)

Chapter 7

Frenkel-Kontorova Model

The Frenkel-Kontorova (FK) model concerns the equilibrium configuration of a one-dimensional chain of masses connected by springs and simultaneously subject to a periodically varying pinning potential. Ignoring the kinetic energy of the mass (since we are only interested in the equilibrium positions) the total potential energy of the springs may be expressed as

$$H_0 = \frac{J}{2}\frac{\pi^2}{b^2}\sum_i (z_{i+1} - z_i - a)^2, \tag{7.1}$$

where z_i is the position of the i-th mass, and a is the "natural" separation between the masses. The substrate potential is of the form

$$V = \frac{K}{2}\sum_i \left(1 - \cos\frac{2\pi z_i}{b}\right), \tag{7.2}$$

with a "natural" period of b. The two competing periodicities, a and b, leads to the complex behavior of the ground state. The sum of the two energies constitutes the FK model, also known as the Frank and Van der Merwe (FvM) model:

$$H = \frac{J}{2}\frac{\pi^2}{b^2}\sum_i (z_{i+1} - z_i - a)^2 + \frac{K}{2}\sum_i \left(1 - \cos\frac{2\pi z_i}{b}\right). \tag{7.3}$$

The difference equation that results from minimizing H, $\partial H/\partial z_i = 0$ reads

$$z_{i+1} + z_{i-1} - 2z_i = \frac{b}{\pi}\frac{K}{J}\sin\left(\frac{2\pi}{b}z_i\right). \tag{7.4}$$

7.1 Nearly identical periods

As a start, we will consider the situation where the two natural periods differ only slightly. It is most convenient to proceed after first making the transformation

$$z_i = ib + (b/\pi)\theta_i, \tag{7.5}$$

which reduces the the FvM model to

$$H = \frac{1}{2}J\sum_i \left(\theta_{i+1} - \theta_i - \pi\frac{a-b}{b}\right)^2 + \frac{K}{2}\sum_i \left(1 - \cos[2\theta_i]\right), \tag{7.6}$$

and the difference equation to

$$\theta_{i+1} + \theta_{i-1} - 2\theta_i = \frac{K}{J}\sin\left(2\theta_i\right). \tag{7.7}$$

As long as $|a-b|/b \equiv \delta \ll 1$, we can regard $\theta_i 1$ to be slowly varying and take the continuum approximation $\theta_{i+1} - \theta_i \approx \partial_x\theta$. In this sense, the discrete lattice position i becomes a continuous variable x and the discrete Hamiltonian becomes a continuum one

$$H = \frac{1}{2}J\int dx\left[\partial_x\theta - \delta\right]^2 + \frac{K}{2}\int dx[1 - \cos(2\theta)]. \tag{7.8}$$

The minimum-energy configuration is the one obeying the differential equation

$$J\partial_x^2\theta - K\sin(2\theta) = 0. \tag{7.9}$$

This particular differential equation is known as the sine-Gordon equation. One can obtain this by taking a variational derivative of H with respect to $\theta(x)$. Alternatively, one can directly work with the discrete equation, Eq. (7.7) and transform it to a differential equation by expand $\theta_{i\pm1} = \theta_i \pm \partial\theta_i/\partial i + (1/2)\partial^2\theta/\partial i^2$.

Note that a circular pendulum under the influence of gravity obeys the equation

$$ml\frac{d^2\theta}{dt^2} = -mg\sin\theta. \tag{7.10}$$

The similarity can be made identical after replacing $\theta \to \theta + \pi/2$ in Eq. (7.9) and treating x as a "time". For this reason the sine-Gordon equation is also known as the "pendulum equation". (It's interesting that seemingly completely different physical phenomena are nevertheless described by the same set of equations.) Like other differential equations appearing in physics, such an equation bears a special function as its solution. In this case, the solution is known as the Jacobi elliptic function.

To see how the Jacobi function arises, first multiply both sides of Eq. (7.9) by $\partial_x\theta$ gives

$$\frac{d}{dx}\Big(J(\partial_x\theta)^2 + K\cos 2\theta\Big) = 0. \tag{7.11}$$

The quantity inside the parenthesis is a constant, which can be parameterized as

$$(\partial_x \theta)^2 + \frac{2K}{J}\cos^2\theta = \frac{2K}{mJ}, \tag{7.12}$$

with the free parameter m. (In fact, this constant quantity is nothing but the total energy in the pendulum problem.) One can multiply both sides by $\sin^2\theta$ and write

$$\left(\frac{d}{dx}\cos\theta\right)^2 = \frac{2K}{mJ}(1-\cos^2\theta)(1-m\cos^2\theta). \tag{7.13}$$

This is exactly the differential equation obeyed by the Jacobi elliptic function $y = \mathrm{sn}(z;\sqrt{m})$

$$\left(\frac{dy}{dz}\right)^2 = (1-y^2)(1-my^2). \tag{7.14}$$

We may thus identify the solution as

$$\cos[\theta(x,m)] = \mathrm{sn}\left(\sqrt{\frac{2K}{mJ}}x, \sqrt{m}\right). \tag{7.15}$$

The sn function reduces to the harmonic sine function when $\sqrt{m} \to 0$: $\mathrm{sn}(z; k \to 0) = \sin z$. For general m, one can use the Mathematica command for a function, **JacobiSN[z,m]**. The Jacobi functions are plotted in Fig. **??** for a selection of m. As m grows, the plot evolves from a sine-like function to a series of soliton-wall type solution. The $m \to 0$ limit should recover the harmonic function, but only by allowing K/J to vanish as m. A vanishing K/J implies that the amplitude of the fluctuation is very small, so one can be justified to expand $\sin(2\theta)$ in Eq. (7.9) to first order. Indeed, one recovers the harmonic equation in this limit.

The evolution of the Jacobi sn function with m has a physical interpretation in the pendulum problem. Taking x to be the time axis, an increasing m describes the pendulum swing that progressively reaches higher and higher heights at the maximum angle. The flattening at the position of the maximum swing means that the pendulum spends more and more time near the turning points before it turns back and begins to swing the other way.

The meaning of $\theta(x)$ as obtained from above is as follows. Going back to the introduction of θ_i in Eq. (7.5), one finds that accumulating the change in θ_i by π amounts to having $N \pm 1$ particles sitting over the substrate periods of $N \times b$. If such an accumulation of phase (hence the particle) occurs over an integer number of potential valleys, the same structure of having $N \pm 1$ particles over N substrate periods will repeat itself. Such structures are "commensurate" because there is a rational value that defines the ratio of particles numbers and the substrate periods over which the particles fit. Otherwise, one says the structure is "incommensurate".

On intuitive grounds, for $a \neq b$ but close to each other, one would expect a 1 : 1 commensurate ground state when $K/J \gg 1$. On the other hand, if $J/K \gg 1$, the inter-particle spacing will be maintained at a' where a is slightly modified from its bare value a due to the interaction with the substrate. Then, there will be a sort of phase transition point at $K/J = (K/J)_c$ separating the incommensurate phase from the commensurate phase. It is the purpose of the subsequent calculation to work out the precise value of the critical strength $(K/J)_c$.

By inserting the solution $\cos[\theta(x)] = \mathrm{sn}(\sqrt{2K/mJ}x, \sqrt{m})$ into the continuum Hamiltonian, Eq. (7.8), one obtains an energy that depends on the parameter m. The determination of m proceeds by the consideration of the minimum energy. We can define the period λ as the length over which $\theta(x)$ increase by 2π. Then one can define the energy density as the energy, Eq. (7.8), evaluated over a period, divided by λ. This quantity, when minimized with respect to m, will be able to fix m.

The period of modulation λ is the value obtained from the integration $\int_0^\lambda [\partial_x \theta] dx = 2\pi$:

$$\begin{aligned} \lambda &= \sqrt{\frac{mJ}{2K}} \int_0^{2\pi} \frac{d\theta}{\sqrt{1 - m\cos^2\theta}} = 4\sqrt{\frac{mJ}{2K}} K_1(m), \\ K_1(m) &= \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - m\cos^2\theta}}. \end{aligned} \tag{7.16}$$

We have introduced the complete elliptic integral of the first kind, $K_1(m)$. The average energy over a period λ picks up contributions from

$$\begin{aligned} &\frac{J}{2\lambda} \int_0^\lambda \left(\frac{\partial\theta}{\partial x}\right)^2 dx = \frac{K}{m}\frac{K_2(m)}{K_1(m)}; \quad K_2(m) = \int_0^{\pi/2} \sqrt{1 - m\cos^2\theta} d\theta, \\ &\frac{J}{2\lambda} \int_0^\lambda [\delta^2 - 2\delta(\partial_x\theta)] dx = \frac{1}{2} J\delta^2 - \pi\delta\sqrt{\frac{JK}{2m}} \frac{1}{K_1(m)}, \\ &-\frac{K}{\lambda} \int_0^\lambda \cos^2\theta dx = \frac{K}{m}\left(\frac{K_2(m)}{K_1(m)} - 1\right). \end{aligned} \tag{7.17}$$

The complete elliptic integral of the second kind, $K_2(m)$, has been introduced. To see how the elliptic integrals $K_1(m)$ and $K_2(m)$ behave, try plotting **EllipticK[m]** and **EllipticE[m]** on Mathematica.

The total average energy density of the solution is

$$E(m) = \frac{K}{m}\left(2\frac{K_2(m)}{K_1(m)} - 1\right) - \pi\delta\sqrt{\frac{JK}{2m}} \frac{1}{K_1(m)}. \tag{7.18}$$

We will make it dimensionless by dividing the energy by K and then defining the quantity $A = J\pi^2\delta^2/2K$:

$$\frac{E(m)}{K} = e(m) = \frac{1}{m}\left(2\frac{K_2(m)}{K_1(m)} - 1\right) - \sqrt{\frac{A}{m}} \frac{1}{K_1(m)}. \tag{7.19}$$

Minimum of $e(m)$ with respect to m fixes the optimal value of m. Using the identities

$$\frac{dK_2}{dm} = \frac{K_2 - K_1}{2m}, \quad \frac{dK_1}{dm} = \frac{K_2 - (1-m)K_1}{2m(1-m)}, \tag{7.20}$$

we can show that

$$\frac{de}{dm} = \frac{K_2(m)[\sqrt{Am} - 2K_2(m)]}{2m^2(1-m)[K_1(m)]^2}. \tag{7.21}$$

The derivative equals zero if

$$2K_2(m) = \sqrt{Am}. \tag{7.22}$$

The elliptic function $2K_2(m)$ varies over π to 2 as m varies from 0 to 1. There is therefore a critical strength of $A \geq A_c \equiv 4$ for which a solution to $de/dm = 0$ exists. Since $A = J\pi^2\delta^2/2K$, this condition amounts to the critical coupling

$$\left(\frac{K}{J}\right)_c = \frac{\pi^2\delta^2}{8}. \tag{7.23}$$

For K/J less than this value, a finite-m, $m < 1$ exists which satisfies Eq. (7.22). $\cos[\theta(x)]$ is a kind of distorted sinusoidal wave depicted in Fig. **??**. The period λ of the ground state configuration can be worked out. With $\lambda = 4\sqrt{\frac{mJ}{2K}}K_1(m)$, and m defined by $\sqrt{Am} = 2K_2(m)$, we have

$$\lambda = \lambda_0 \left(\frac{2}{\pi}\right)^2 K_1(m)K_2(m), \tag{7.24}$$

where the natural period $\lambda_0 = 2\pi/\delta$ has been introduced. $\theta(x)$ increases by 2π over this length, however this length is generally incommensurate with respect to the substrate potential. We are thus in the "incommensurate phase". The period λ diverges at $m = 1$ (since $K_1(1) = \infty$), then decreases logarithmically as $\lambda/\lambda_0 \sim -\log(1-m)$ for $m < 1$. The energy density for the incommensurate phase is, after some algebra,

$$e(m) = -\frac{1}{m}. \tag{7.25}$$

If Eq. (7.22) is not fulfilled, it implies that $de/dm < 0$ for all values of m, and thus the minimum energy occurs at $m = 1$. When $m = 1$, as shown in Fig. **??**, the solution is a solitonic one where the angle jumps suddenly from 0 to π and back. In particular $\theta = 0$ means that the angle θ remains "pinned" at that value regardless of x, or that z_i remains pinned at $z_i = b \times i$. This describes the situation where each well of the potential contains one mass. The pinning phase, or the commensurate phase as it is more commonly called, occurs when

the strength of the pinning potential K/J is larger than $(K/J)_c$. The critical coupling in turn grows with δ^2, the degree of incommensurability between the substrate and the spring interaction.

To better understand the evolution of the solution with the ratio $J\delta^2/K$, we re-write Eq. (7.15) in the form

$$\cos\theta = \mathrm{sn}\left(\frac{\pi\delta}{2K_2(m)}x;\sqrt{m}\right). \tag{7.26}$$

For m close to but less than one, the sn function is approximated by the hyperbolic function as

$$\cos\theta = \mathrm{sn}\left(\frac{\pi\delta}{2K_2(m)}x;\sqrt{m}\right) \approx \tanh(\delta x). \tag{7.27}$$

The corresponding angle θ varies from π to 0 as x passes through $x = 0$. This is the single-soliton limit with one domain wall separating $\theta = 0$ from $\theta = \pi$ regions. For sufficiently small δ or a sufficiently large K/J this is the stable solution. As m decreases, there occurs more and more periodic modulation of $\cos\theta$ until the period of the modulation agrees with the free value $\lambda_0 = 2\pi/\delta$. The best way to see this is by plotting the sn function using the **JacobiSN[z,m]** function on Mathematica while gradually reducing m from 1.

7.2 Commensurate periods

The previous discussion has been restricted to the case of nearly identical periodicities, $a \approx b$. We saw that there existed a critical ratio of K/J that scales with δ^2, beyond which a 1 : 1 commensurate state is the ground state. An incommensurate phase is achieved for K/J less than the critical value. If $a = b$, the critical coupling is zero, and one is always in the commensurate phase.

In this section, we consider the case where a and b are commensurate: $a/b = p/q < 1$ with p and q relative primes. As earlier, we find it convenient to first transform to a different variable,

$$z_i = a \times i + \frac{b}{\pi}\phi_i = \frac{p}{q}bi + \frac{b}{\pi}\phi_i. \tag{7.28}$$

The new variable ϕ_i is subject to the energy functional

$$H = \frac{1}{2}J\sum_i (\phi_{i+1} - \phi_i)^2 + \frac{K}{2}\sum_i \left(1 - \cos[2\pi i\frac{p}{q} + 2\phi_i]\right), \tag{7.29}$$

and the equation of motion

$$J(\phi_{i+1} + \phi_{i-1} - 2\phi_i) = K\sin\left(2\pi i\frac{p}{q} + 2\phi_i\right). \tag{7.30}$$

The equation is invariant under the translation of the lattice spacing by q units: $i \to i+q$. A solution that obeys such an equation also has the periodic property, $\phi_{i+q} = \phi_i$, hence can be Fourier-expanded in integer harmonics of p/q.

$$\phi_i = \sum_n t_n \sin\left(2\pi\frac{p}{q}ni + 2u_n\right). \tag{7.31}$$

Assuming that the first harmonic (p/q) is the most dominant part of the Fourier expansion, we can write down the solution ϕ_i in the general form

$$\phi_i = t\sin\left(2\pi\frac{p}{q}i + 2u\right) + v. \tag{7.32}$$

Here t, u and v are constants to be determined from the minimization of the energy. Substituting into the kinetic term and averaging over q lattice spacings gives

$$\begin{aligned}&\frac{1}{q}\frac{J}{2}\sum_{i=i_0+1}^{i_0+q}(\phi_{i+1}-\phi_i)^2\\ &= Jt^2\sin^2\left(\pi\frac{p}{q}\right)\frac{1}{q}\sum_i\left(1+\cos[4\pi\frac{p}{q}i + 2\pi\frac{p}{q} + 4u]\right) = Jt^2\sin^2\left(\pi\frac{p}{q}\right).\end{aligned} \tag{7.33}$$

Substituting into the potential term and averaging over q lattice spacings gives

$$\begin{aligned}&\frac{K}{2} - \frac{K}{2}\frac{1}{q}\sum_{i=i_0+1}^{i_0+q}\cos\left[2\pi\frac{p}{q}i + 2t\sin\left(2\pi\frac{p}{q}i+2u\right)+2v\right] =\\ &\frac{K}{2} - \frac{K}{2}\frac{1}{q}\sum_{i=i_0+1}^{i_0+q}\cos\left[2\pi\frac{p}{q}i+2v\right]\cos\left[2t\sin\left(2\pi\frac{p}{q}i+2u\right)\right]\\ &+\frac{K}{2}\frac{1}{q}\sum_{i=i_0+1}^{i_0+q}\sin\left[2\pi\frac{p}{q}i+2v\right]\sin\left[2t\sin\left(2\pi\frac{p}{q}i+2u\right)\right].\end{aligned} \tag{7.34}$$

At this point we invoke the identity

$$\begin{aligned}\cos[2t\sin\left(2\pi\frac{p}{q}i+2u\right)] &= J_0(2t) + 2\sum_{s=1}^{\infty}J_{2s}(2t)\cos[2s\left(2\pi\frac{p}{q}i+2u\right)]\\ \sin[2t\sin\left(2\pi\frac{p}{q}i+2u\right)] &= 2\sum_{s=0}^{\infty}J_{2s+1}(2t)\sin[(2s+1)\left(2\pi\frac{p}{q}i+2u\right)].\end{aligned} \tag{7.35}$$

There can be two non-vanishing terms in the potential energy. The $2J_1(2t)\sin[2\pi(p/q)i+2u]$ from the second line of Eq. (7.35) contributes $(K/2)J_1(2t)\cos[2(u-v)]$ to the potential energy. A second contribution to the potential energy comes from either $2s = q-1$ (first line of Eq. (7.35)) or $2s+1 = q-1$ (second line of Eq. (7.35)), depending on whether q is odd or even. In either case, one has $-(K/2)J_{q-1}(2t)\cos[2(v-u)+2qu]$. All told, the averaged potential energy becomes

$$\frac{K}{2} + \frac{K}{2} J_1(2t) \cos[2(u-v)] - \frac{K}{2} J_{q-1}(2t) \cos\left[2(v-u) + 2qu\right]. \tag{7.36}$$

Since the kinetic energy does not depend on u or v, the optimal u, v are solely governed by the potential energy. Assuming that we are considering the small t and hence $J_1(2t) \gg J_{q-1}(2t)$, it comes out that $2(u-v) = \pi$ from the $J_1(2t)$ term of the potential energy. From the $J_{q-1}(2t)$ term, it then follows that $u = \pi/2q$. Inserting these back, the total averaged energy becomes

$$E = Jt^2 \sin^2\left(\pi \frac{p}{q}\right) + \frac{K}{2}[1 - J_1(2t)] - \frac{K}{2} J_{q-1}(2t). \tag{7.37}$$

For small t and $q \geq 3$, one can replace $J_1(2t) \approx t$ and find the minimum energy condition

$$t = \frac{K}{4J \sin^2(\pi p/q)}, \tag{7.38}$$

and the ground state ϕ_i given by

$$\phi_i = \frac{K}{4J \sin^2(\pi p/q)} \sin\left(2\pi \frac{p}{q} i + \frac{\pi}{q}\right) + \frac{\pi}{2q} - \frac{\pi}{2}. \pmod{(\pi)} \tag{7.39}$$

7.3 Nearly commensurate periods

Finally, we move to the case where a/b is nearly commensurate, $a/b = p/q + \delta/\pi$. Then applying the same transformation, Eq. (7.28), gives FvM model

$$H = \frac{1}{2} J \sum_i \left(\phi_{i+1} - \phi_i - \delta\right)^2 + \frac{K}{2} \sum_i \left(1 - \cos[2\pi i \frac{p}{q} + 2\phi_i]\right). \tag{7.40}$$

In this case we assume a more general ansatz for ϕ_i that allows for spatial variations of u and v:

$$\phi_i = t \sin\left(2\pi \frac{p}{q} i + 2u_i\right) + v_i. \tag{7.41}$$

We expect that the averaged potential energy for slowly varying u_i and v_i can still be written in the form

$$\sum_i \left(\frac{K}{2} + \frac{K}{2} J_1(2t) \cos[2(u_i - v_i)] - \frac{K}{2} J_{q-1}(2t) \cos\left[2(v_i - u_i) + 2qu_i\right]\right). \tag{7.42}$$

The kinetic energy takes on the approximate form

$$Jt^2 \sum_i \sin^2\left(\pi \frac{p}{q} + u_{i+1} - u_i\right) + \frac{J}{2} \sum_i (v_{i+1} - v_i - \delta)^2. \tag{7.43}$$

As before, we may set $u_i = v_i + \pi/2$ to minimize the $J_1(2t)$ term in the potential energy, so the total energy will read

$$Jt^2 \sum_i \sin^2\left(\pi\frac{p}{q} + u_{i+1} - u_i\right) + \frac{J}{2}\sum_i (u_{i+1} - u_i - \delta)^2$$
$$\sum_i \left(-\frac{K}{2}J_1(2t) + \frac{K}{2}J_{q-1}(2t)\cos\left[2qu_i\right]\right) \tag{7.44}$$

Minimization still requires that t be given by Eq. (7.38). We can shift u_i by $\pi/2q$, to get the energy in terms of the slowly varying u_i:

$$Jt^2 \sum_i \sin^2\left(\pi\frac{p}{q} + u_{i+1} - u_i\right) + \frac{J}{2}\sum_i (u_{i+1} - u_i - \delta)^2 - \frac{K}{2}J_{q-1}(2t)\sum_i \cos\left[2qu_i\right]$$
$$\approx \frac{J}{2}\sum_i (u_{i+1} - u_i - \delta)^2 - \frac{K}{2}J_{q-1}(2t)\sum_i \cos\left[2qu_i\right]. \tag{7.45}$$

The second line follows because of the smallness of the first term in the kinetic energy compared to the second by a factor $t^2 \sim (K/J)^2$. We have arrived at the KdV equation with the potential energy scale renormalized by $K \to KJ_{q-1}(2t) \sim K(K/J)^{q-1}$. The commensurate-incommensurate transition is expected to take place at $K_r/J = \pi^2\delta^2/8$, $K_r \approx K(K/J)^{q-1}$.

7.4 Finals Project

Consider reducing the ratio a/b from one down to a small value. As a/b decreases below a threshold value, a transition of the ground state from a $1:1$ commensurate phase to an incommensurate phase will occur according to the discussion of Sec. 7.1. As a/b passes through the value $1/2$, another commensurate phase ($1:2$ phase) forms centered at $a/b = 1/2$, but the width of this commensurate phase in terms of a/b will be narrower than the corresponding width of the $1:1$ phase according to the discussion of Sec. 7.3. In fact, a commensurate phase will be formed for every rational number $a/b = p/q$ with varying widths of stability. The smaller-q commensurate phases will be more easily formed due to a larger width of stability. Using the technique of simulated annealing you've learned in class, come up with the phase diagram of FK model with one axis being the a/b ratio (varying between 0 and 1), and the other axis being K/J.

Chapter 8

Problems

8.1 Oscillation

1. *Forced damped oscillator*:

   (a) Solve the forced oscillation of a damped harmonic oscillator

   $$\frac{d^2x}{dt^2} + 2\lambda\frac{dx}{dt} + \omega_0^2 x = \frac{f}{m}\cos\gamma t, \tag{8.1}$$

   using the complex expression $x = Be^{i\gamma t}$.

   (b) From the solution, work out the amount of energy dissipated per unit time through friction, dE/dt.

2. A pendulum of length l and mass m is suspended at a point of support. The point of support itself is moving vertically according to $y(t) = a\cos[\gamma t]$. The Lagrangian which describe the motion of the pendulum is

   $$L = \frac{1}{2m}ml^2\dot{\phi}^2 - mal\gamma^2\cos\phi\cos[\gamma t] + mgl\cos\phi. \tag{8.2}$$

   (a) Write down the equation of motion.

   (b) When γ is much greater than the natural frequency of the pendulum, one can introduce an effective potential which is given by

   $$U_{eff}(\phi) = -mgl\cos\phi + \frac{ma^2\gamma^2}{4}\sin^2\phi. \tag{8.3}$$

   Under what condition do we obtain a stable equilibrium at $\phi = \pi$? Compare the curvature of the potential expanded up to second order in the angle around $\phi = 0$ and $\phi = \pi$. Discuss which angular position would allow a more stable oscillation.

3. A mass m with coordinate x is moving according to the Lagrangian

$$L = \frac{1}{2}m\left(\frac{dx}{dt}\right)^2 - \frac{1}{2}m\omega^2x^2 - \frac{1}{3}m\alpha x^3 - \frac{1}{4}m\beta x^4. \tag{8.4}$$

(a) Write down the equation of motion. Describe, in general terms, the new features of the solution that will appear as a result of the nonlinear effects.

(b) Work out the solution in series form $x = x^{(1)} + x^{(2)}$, taking $x^{(1)} = a\cos\omega_0 t$.

4. We want to solve the problem

$$\frac{d^2x}{dt^2} + \omega^2(t)x = 0, \tag{8.5}$$

where $\omega^2(t) = \omega_0^2[1 + h\cos\gamma t]$ and furthermore γ is close to twice the natural frequency of the harmonic oscillator. Then one can set $\gamma = 2\omega_0 + 2\varepsilon$ with ε being a small number. Find a solution of the problem of the form

$$x(t) = a(t)\cos[(\omega_0 + \varepsilon)t] + b(t)\sin[(\omega_0 + \varepsilon)t]. \tag{8.6}$$

Work out the equations obeyed by $a(t)$ and $b(t)$, ignoring terms of higher order in h or ε. Solve for $a(t)$ and $b(t)$ and show that one of the solutions must diverge exponentially with time.

5. In class we learned that if a particle motion is subject to both a slow and a fast force,

$$m\frac{d^2x}{dt^2} = -\frac{dU}{dx} + f(x,t), \tag{8.7}$$

then one can incorporate the effects of the fast force $f(x,t)$ by introducing the notion of an effective potential. If $f(x,t)$ is of the form $f(x,t) = f(x)\cos\gamma t$, we were able to show that the effective potential is of the form

$$U_{eff}(X) = U(X) + \frac{1}{4m\gamma^2}f(X)^2, \tag{8.8}$$

written exclusively in terms of the "slow" variable X. Using these results work on the following problem.

(a) Derive the Lagrangian of a pendulum of length l and mass m, whose point of support is moving vertically according to $y = a\cos\gamma t$.

(b) Write down the corresponding equation of motion, and identify the fast and slow parts of the force from it.

(c) Using the above-mentioned result, write down the effective potential for the pendulum. Work out the equilibrium position for the angle ϕ when γ is very small and when it is very large.

6. *Coupled harmonic oscillator*:

   (a) Work out the equation of motion of a coupled harmonic oscillator whose Lagrangian is given by

   $$L = \frac{1}{2}\left(\dot{x}^2 + \dot{y}^2\right) - \frac{1}{2}\omega_0^2(x^2 + y^2) + \alpha xy. \tag{8.9}$$

   (b) Solve for x and y. What are the normal coordinates? Prove that the Lagrangian re-written in terms of the normal coordinates indeed become diagonal, i.e. no coupling term arises between the coordinates.

7. Solve the problem of the chain of coupled harmonic oscillator of equal force constant k and alternating masses $m_1 = m$, $m_2 = 2m$, and $m_3 = 3m$. Calculate the three eigenfrequencies at the zero wave number. Show that only one mode occurs at zero energy.

8. Motion of a spherical pendulum is governed by the Lagrangian

   $$L = \frac{1}{2}(\dot{\theta}^2 + \dot{\phi}^2 \sin^2\theta) + \cos\theta. \tag{8.10}$$

   Show that the equation of motion can be expressed as $\ddot{\theta} = -\frac{\partial V_{eff}(\theta)}{\partial \theta}$. What is the effective potential $V_{eff}(\theta)$? Write down the conserved energy.

8.2 Others

1. A particle of mass m moving with velocity $\mathbf{v}_1$ leaves a half-space in which its potential energy is a constant U_1 and enters another in which its potential energy is a different constant U_2. Determine the change in the direction of motion of the particle. Express your answer in terms of $\sin\theta_1/\sin\theta_2$, where θ_1 and θ_2 are the angles between the normal of the plane and the velocities $\mathbf{v}_1$ and $\mathbf{v}_2$, respectively.

2. Determine the period of oscillation, as a function of energy, when a particle of mass m moves in a field of the potential energy given by (a) $U(x) = kx^2$, (b) $U(x) = kx^4$, (c) $U(x) = U_0 \tan^2 \alpha x$.

3. The Lagrangian concerning the motion of two particles which interact with the potential energy that depends only on the distance between them can be written as

   $$L = \frac{1}{2}m_1\mathbf{v}_1^2 + \frac{1}{2}m_2\mathbf{v}_2^2 - U(|\mathbf{r}_1 - \mathbf{r}_2|). \tag{8.11}$$

   (a) Simplify the Lagrangian by passing to the center-of-mass coordinates which satisfy $m_1\mathbf{r}_1 + m_2\mathbf{r}_2 = 0$. Express the Lagrangian using the relative coordinate $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$. What is the reduced mass m in terms of m_1 and

m_2?

(b) Due to the conservation of angular momentum vector, the motion of $\mathbf{r}$ is confined to a plane. Express the reduced Lagrangian obtained in (a) in terms of the cylindrical coordinates (r, ϕ). How is the conserved angular momentum M_z expressed in terms of (r, ϕ)? The motion is assumed to take place in the $x - y$ plane.

(c) Using the result of (b), express the reduced Lagrangian obtained in (a) in terms of r and its time derivative dr/dt alone. What is the conserved energy in terms of r and dr/dt?

(d) Express the time t in terms of the integral of the function that depends on r. Express the angle ϕ in terms of the integral of another function that depends on r.

4. Suppose the potential energy satisfies the homogeneity condition

$$U(\alpha\mathbf{r}_1, \alpha\mathbf{r}_2, \cdots, \alpha\mathbf{r}_n) = \alpha^k U(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_n). \tag{8.12}$$

(a) What can you say about the relation of the time scales of a similar motion that take place on the length scales of l and l', respectively. Denote the time scales as t and t'.

(b) Deduce the relation between the period of oscillation T and the typical size of the orbit L if the potential energy is of the type $U(x) \sim x^k$, $k > 0$.

5. Which components of momentum $\mathbf{P}$ and angular momentum $\mathbf{M}$ are conserved in motion in the following fields? (a) the field of an infinite homogeneous plane, (b) that of an infinite homogeneous cylinder, (c) that of two points, (d) that of a homogeneous cone. Explain your answer carefully.

6. Starting from the definition of the Hamiltonian H

$$H = \sum_i p_i \dot{q}_i - L, \tag{8.13}$$

where L is the Lagrangian of the mechanical system, derive the Hamilton's equation of motion

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}. \tag{8.14}$$

You can do this by working out the small variations dH for the corresponding Hamiltonian H. The generalized momentum p_i is defined from the Lagrangian as $p_i = \partial L/\partial \dot{q}_i$.

7. Using the Poisson bracket relations between the coordinates and the momenta

$$[p_i, q_j] = \delta_{ij}, \tag{8.15}$$

derive the Poisson bracket for the x and y components of the angular momentum vector, $[M_x, M_y]$.

8. In class we considered the notion of the adiabatic invariant. For the harmonic oscillator with a time-dependent spring constant whose Hamiltonian reads

$$H = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\omega(t)^2x^2, \tag{8.16}$$

such an adiabatic invariant turned out to be $I = E(t)/\omega(t)$. Show by an explicit calculation that dI/dt averaged over a single period T is indeed zero provided we could regard the time dependence of $\omega(t)$ to be slow, i.e. $T|d\omega/dt| \ll \omega$.

9. The Frenkel-Kontorova model is defined as the classical model given by the following energy functional

$$H = \frac{J}{2}\frac{\pi^2}{b^2}\sum_i (z_{i+1} - z_i - a)^2 + \frac{K}{2}\sum_i \left(1 - \cos\frac{2\pi z_i}{b}\right). \tag{8.17}$$

z_i represents the one-dimensional position of the i-th "ball" that interact with each other through the first term of H, and simultaneously subject to the substrate potential due to the second term of H given above.

(a) Derive the equation satisfied by z_i which would minimize the total energy.

(b) Assume that the two competing periods a and b are sufficiently close that one can take $(a-b)/b$ to be very small. In this case one can write z_i as $z_i = i \times b + (b/\pi)\theta_i$ and substitute into the equation derived in (a). Argue that the new variable θ_i may be regarded as "slowly varying" in space and that one can take a continuum approximation for it. What is the differential equation obeyed by $\theta(x)$, where x is the continuum representation of the balls' position.

10. The kinetic energy of a rigid body is expressed as $T = \frac{1}{2}\omega_a I_{ab}\omega_b$ where the inertia tensor I_{ab} is given by

$$I_{ab} = \sum_i m_i\Big(\mathbf{r}_i\cdot\mathbf{r}_i\delta_{ab} - (\mathbf{r}_i)_a(\mathbf{r}_i)_b\Big) = \int \rho(\mathbf{r})\Big(\mathbf{r}\cdot\mathbf{r}\delta_{ab} - (\mathbf{r})_a(\mathbf{r})_b\Big). \tag{8.18}$$

Describe the process of diagonalizing the inertia tensor and using the eigenvectors and eigenvalues to re-write the kinetic energy in diagonal form.

11. A point on a rigid body executes a motion according to the first-order differential equation

$$\dot{\mathbf{r}} = \boldsymbol{\omega} \times \mathbf{r}, \tag{8.19}$$

where the rotation vector $\boldsymbol{\omega}$ itself is a time-dependent function. Write down the solution of such an equation, and explain your answer.

12. Consider a charged particle moving in two dimensions, subject to the Lagrangian

$$L = \frac{m}{2}\dot{\mathbf{r}}^2 - e(A_0 - \dot{\mathbf{r}} \cdot \mathbf{A}). \tag{8.20}$$

Here A_0 and $\mathbf{A}$ are the scalar and vector potentials, respectively. Derive the Lagrange's equation of motion and show that it is equivalent to Lorentz's force law.

13. A system of two particles obey the Lagrangian

$$L = x_1\dot{y}_1 + x_2\dot{y}_2 - V(|\mathbf{r}_1 - \mathbf{r}_2|). \tag{8.21}$$

Here $\mathbf{r}_1 = (x_1, y_1)$ and $\mathbf{r}_2 = (x_2, y_2)$ are the coordinates of the two particles in two dimensions, $V(|\mathbf{r}_1 - \mathbf{r}_2|)$ is a central potential depending on the relative distance of the two particles. Derive the Lagrange's equation and discuss the nature of the relative motion. Note that the kinetic energy part is not the usual one.