1 Basic Mechanics

1.1 Newtonian mechanics a brief review

Momentum and Center of Mass

- Newton’s equations of motion for a system of particles reads
  \[ \frac{dp_a}{dt} = F_a \]  
  where \( a = 1 \ldots N \) labels the particles. Here \( p_a = m_a v_a \). We usually divide up the forces on the \( a \)-the particle into external forces acting on the system from outside, and internal forces acting between pairs of particles:
  \[ F_a = F_a^{\text{ext}} + \sum_{b \neq a} F_{ab} \]
  external forces internal forces
  
  Here
  \[ F_{ab} = \text{Force on particle } a \text{ by } b, \]
  and of course we have Newton’s equal and opposite rule
  \[ F_{ab} = -F_{ba}. \]

- Summing over the particles we find (after using Eq. (1.4)) that the internal forces cancel and the total change in momentum per time is the sum of external forces
  \[ \frac{dP_{\text{tot}}}{dt} = F_{\text{tot}}^{\text{ext}} \]
  where \( P_{\text{tot}} = \sum_a p_a \) and \( F_{\text{tot}}^{\text{ext}} = \sum_a F_{a}^{\text{ext}} \). If there are no external forces then \( P_{\text{tot}} \) is constant

- The velocity of the center of mass is
  \[ v_{\text{cm}} = \frac{P_{\text{tot}}}{M_{\text{tot}}} = \frac{1}{M_{\text{tot}}} \sum_a m_a v_a. \]

  The position of the center of mass (relative to an origin \( O \)) is
  \[ R_{\text{cm}} = \frac{1}{M_{\text{tot}}} \sum_a m_a r_a. \]

Angular momentum:

- Angular momentum is defined with respect to a specific origin \( O \) (i.e. \( r_a \) depends on \( O \)) which is not normally notated
  \[ \ell_{a,O} \equiv \ell_a \equiv r_a \times p_a. \]
It evolves as
\[
\frac{d\ell}{dt} = r_a \times F_a \tag{1.9}
\]

- The total angular momentum \( L_{\text{tot}} = \sum_a \ell_a \) changes due to the total external torque
\[
\frac{dL_{\text{tot}}}{dt} = \tau_{\text{ext}}^{\text{tot}} , \tag{1.10}
\]

where \( \tau_{\text{ext}}^{\text{tot}} = \sum_a r_a \times F_a^{\text{ext}} \) were we have generally assumed that the internal forces are radially directed \( F_{ab} \propto (r_a - r_b) \)

- The angular momentum depends on the origin \( O \). Writing the position of the particle relative to the center of mass as \( \Delta r_a \), i.e.
\[
\ell_a = R_{\text{cm}} + \Delta r_a , \tag{1.11}
\]

the angular momentum of the system about \( O \) is
\[
L_O = R_{\text{cm}} \times P_{\text{tot}} \quad + \quad \sum_a \Delta r_a \times p_a \quad \text{Ang-mom of center of mass about } O \quad \text{Ang-mom about the cm} \tag{1.12}
\]

Energy

- Energy conservation is derived by taking the dot product of \( v \) with \( dp/\text{dt} \). We find that the change in kinetic energy (on the \( a \)-the particle) equals the work done (on the \( a \)-particle).
\[
\frac{1}{2} m_a v_a^2(t) \bigg|_{t_2}^{t_1} = W_a \tag{1.13}
\]

where the work is
\[
W_a = \int_{r_a(t_1)}^{r_a(t_2)} F_a \cdot dr_a \tag{1.14}
\]

- Potential Energy. For conservative forces the force can be written as (minus) the gradient of a scalar function which we call the potential energy
\[
F_a = -\nabla r_a U \tag{1.15}
\]

Consider the potential energy \( U_{12} \) between particle 1 and 2. Since the force is equal and opposite
\[
F_{12} = -\nabla r_1 U_{12}(r_1, r_2) = +\nabla r_2 U_{12}(r_1, r_2) = -F_{21} \tag{1.16}
\]

and this is used to conclude that interaction potential between two particles is of the form
\[
U_{12}^{\text{int}} = U(|r_1 - r_2|) \tag{1.17}
\]

Typically we divide up the potential into an external potential and the internal ones
\[
U(r_a) = U^{\text{ext}}(r_a) + \frac{1}{2} \sum_{ab, a \neq b} U_{ab}^{\text{int}}(r_a, r_b) \tag{1.18}
\]

The sum over the internal potentials comes with a factor of a half because the energy between particle-1 and particle-2 is counted twice in the sum, e.g. for just two particles
\[
U_{12}^{\text{int}}(r_1, r_2) = \frac{1}{2} \left( U(|r_1 - r_2|) + U(|r_2 - r_1|) \right). \tag{1.19}
\]
1.1. NEWTONIAN MECHANICS A BRIEF REVIEW

• Energy. The total energy is

\[ E = \sum_a \frac{1}{2} m_a v_a^2 + U^{\text{ext}}(r_a) + \frac{1}{2} \sum_{ab, a \neq b} U^{\text{int}}_{ab}(r_a, r_b) \]  

(1.20)

and is constant if there are no non-conservative forces. If there are non-conservative forces then

\[ E(t_2) - E(t_1) = W_{NC} \]  

(1.21)

where the work done by the non-conservative forces is \( W_{NC} = \sum_a \int F^{NC}_a \cdot dr_a \)

• It is convenient to measure velocities relative to the center of mass

\[ v_a = v_{cm} + \Delta v_a \]  

(1.22)

where \( \Delta v_a = \dot{\Delta} r_a \), then the kinetic energy

\[ K = \frac{1}{2} M_{tot} v_{cm}^2 + \sum_a \frac{1}{2} m_a \Delta v_a^2 \]  

(1.23)

KE of center-mass  KE relative to center-mass

Galilean invariance:

• Consider newton’s laws then for an isolated system of particles

\[ \frac{dp_a}{dt} = F_a \]  

(1.24)

where \( F_a = -\nabla r_a U \) with

\[ U = \frac{1}{2} \sum_{ab, a \neq b} U^{\text{int}}_{ab}(|r_a - r_b|) \]  

(1.25)

Here the space-time coordinates are measured by an observer \( O \) with origin.

Then consider an observer \( O' \) moving with constant velocity \(-u\) relative to \( O \). The “new” coordinates (those measured by \( O' \)) are related to the old coordinates via a Galilean boost

\[ r_a \rightarrow r'_a = r_a + ut \]  

(1.26)

\[ t \rightarrow t' = t \]  

(1.27)

The potential which only depends on \( r_a - r_b \) is independent of the shift. The observer measures

\[ v_a \rightarrow v'_a = v_a + u \]  

(1.28)

\[ p_a \rightarrow p'_a = p_a + m_a u \]  

(1.29)

The equations of motion for observer \( O' \) are unchanged

\[ \frac{dp'_a}{dt'} = F'_a \quad F' \equiv \nabla_{r'} U(|r'_a - r'_b|) \]  

(1.30)
1.2 The action and the Euler Lagrange equations

- The action
  \[ S[r(t)] = \int_{t_1}^{t_2} dt L(r, \dot{r}, t) \]  
  takes an arbitrary path \( r(t) \) (which may not satisfy the EOM) and returns a number. It is called a functional.

- The action principle says that the path \( r(t) \) that satisfies the EOM (sometimes called the classical or “on-shell” path) is an extremum the action\(^1\). This means that if we replace the on-shell path \( r(t) \) with \( r(t) \to r(t) + \delta r(t) \) for an arbitrary (small) function \( \delta r(t) \) that vanishes near \( t_1 \) and \( t_2 \) then the action is unchanged
  \[ S[r(t) + \delta r(t)] = S[r(t)] \] when \( r(t) \) is “on-shell”, i.e. satisfies the EOM

- Generally we define
  \[ \delta S[r(t), \delta r(t)] \equiv S[r(t) + \delta r(t)] - S[r(t)] \] and note that \( \delta S[r, \delta r] \) depends on both the path and the variation. The requirement that \( \delta S = 0 \) determines the equation of motion. You should be able to prove that when \( \delta S = 0 \) for an arbitrary variation, the equations of motion are (in 1d)
  \[ \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} = \frac{\partial L}{\partial x} \]  
  we call
  \[ p_A = \frac{\partial L}{\partial \dot{q}^A} \] the canonical momentum conjugate to \( q^A \) \( A = 1 \ldots N \)
  \[ F_A = \frac{\partial L}{\partial q^A} \] the generalized force associated with \( q^A \)

- If a coordinate \( q^A \) does not appear in the Lagrangian (but of course \( \dot{q}^A \) does or it wouldn’t appear at all), the variable is called cyclic. For a cyclic coordinate we have from the Euler Lagrange equations (Eq. (1.37))
  \[ \frac{dp_A}{dt} = 0 \] i.e. \( p_A \) is a constant of the motion.

The hamiltonian function

- The hamiltonian (or energy) function (sometimes called the “first integral”) is
  \[ h(q, \dot{q}, t) = p\dot{q} - L(q, \dot{q}, t) = \frac{\partial L}{\partial \dot{q}} \dot{q} - L(q, \dot{q}, t) \] and obey the equation of motion
  \[ \frac{dh}{dt} = -\frac{\partial L}{\partial t} . \]

\( h(q, \dot{q}, t) \) is therefore constant if \( L \) does not depend explicitly on time.

\(^1\)Sometimes for clarity we will put a bar, e.g., \( \bar{r}(t) \) to indicate that this path is on-shell, i.e. that it satisfies the EOM.
1.2. THE ACTION AND THE EULER LAGRANGE EQUATIONS

- If more then one coordinate is involved then
  \[ h(q^A, \dot{q}^A, t) = \sum_A p_A q^A - L \]  
  \[ = \frac{\partial L}{\partial \dot{q}^A} \dot{q}^A - L \]  
  \[ (1.43) \]
  \[ (1.44) \]
  where we have and will from now on follow the summation convention, where repeated indices are summed over.

- We will distinguish the hamiltonian function \( h(q, \dot{q}, t) \), which is a function of \( q, \dot{q} \), and \( t \), from the Hamiltonian \( H(p, q, t) \) which is a function of \( q \) and \( p \) and \( t \) through the Legendre transform (more later). Thus \( p_A(q, \dot{q}, t) \) in the hamiltonian function (Eq. (1.43)) is a function of the \( q \) and the \( \dot{q} \), while in the Hamiltonian the \( \dot{q} \) is a function of \( q \) and \( p \).

- For a rather general Lagrangian
  \[ L = \frac{1}{2} a_{ij}(q) \dot{q}^i \dot{q}^j + b_i(q) \dot{q}^i - U(q), \]  
  \[ (1.45) \]
  (which is the form of the Lagrangian for a particle in a magnetic field or gravity) the hamiltonian function is
  \[ h(\dot{q}, q, t) = \frac{1}{2} a_{ij}(q) \dot{q}^i \dot{q}^j + U(q) \]  
  \[ (1.46) \]
  The fact that the hamiltonian function is independent of \( b_i \) is closely related to the fact that magnetic fields do no work.

The period of one dimensional motion

- For one dimensional Lagrangian’s of the form
  \[ L = \frac{1}{2} m(q) \dot{q}^2 - V_{\text{eff}}(q) \]  
  \[ (1.47) \]
  The first integral is
  \[ E = \frac{1}{2} m(q) \dot{q}^2 + V_{\text{eff}}(q) \]  
  \[ (1.48) \]
  You should be able to show that the this first integral equation can be used to determine \( q(t) \) implicitly. Integrating from \((t_0, q_0)\) to \((t, q(t))\) yields
  \[ \pm \int_{q_0}^{q(t)} dq \left( \frac{m(q)}{2(E - V_{\text{eff}}(q))} \right)^{1/2} = t - t_0, \]  
  \[ (1.49) \]
  which, when inverted, gives \( q(t) \). The plus sign is when \( q \) is increasing in time, while the minus sign is when \( q(t) \) is decreasing in time.

- In a typical case the potential \( V_{\text{eff}}(q) \) and energy \( E \) is shown below
For the specified energy, the motion is unbounded for $q > q_c$, and oscillates between when $q_A < q < q_B$. $q_A$, $q_B$ and $q_C$ are called turning points. The period $T(E)$ is the time it takes to go from $q_A$ to $q_B$ and back. Thus half a period $T(E)/2$ is the time it takes to go from $q_A$ to $q_B$ or

$$T(E)/2 = \int_{q_A}^{q_B} dq \left( \frac{m(q)}{2(E - V_{\text{eff}}(q))} \right)^{1/2}.$$  (1.50)
1.3 The Hamiltonian Formalism, the Routhian, and the Legendre Transform

The Hamiltonian formalism: basic version

- Let the Lagrangian be a convex function of the velocity \(v_q \equiv \dot{q}\). In one dimension this means that the momentum \(p = \partial L / \partial v_q\) is an increasing function of the velocity \(v_q \equiv \dot{q}\), i.e. \(\partial^2 L / \partial v_q^2 > 0\). This means there is one value of the velocity for given momentum \(p\), \(\dot{q}(p)\). Clearly \(L \propto v^2\) is convex.

In higher dimensions we require that \(\partial^2 L / \partial \dot{q}_i \partial \dot{q}_j\) is a positive definite matrix. This means that for a given value of \(p_i\) there is a unique value of the velocity vector \(v_i = \dot{q}_i(p)\) at fixed \(q\).

- With convex function \(L(q)\) a Legendre transform useful, and trades the velocity dependence of the Lagrangian dependence for the momentum dependence \(p\) of the Hamiltonian.

First note

\[
dL = pd\dot{q} + \frac{\partial L}{\partial q} dq + \frac{\partial L}{\partial t} dt
\]

\text{“spectators”} \hspace{1cm} (1.51)

We can trade the \(d\dot{q}\) for \(dp\) by looking at \(L - p\dot{q}\), or, as is conventional, minus this quantity. Thus we define

\[
H(p, q, t) = p \dot{q}(p) - L(\dot{q}(p), q, t)
\]

where \(\dot{q}(p)\) is determined from \(p\) at fixed \(q\) and \(t\), i.e. we must invert the relation

\[
p = \frac{\partial L(\dot{q}, q, t)}{\partial \dot{q}} \Rightarrow \text{determines } \dot{q}(p)
\]

We have (do it yourself!)

\[
dH(p, q, t) = \dot{q} dp - \left(\frac{\partial L}{\partial q} dq + \frac{\partial L}{\partial t} dt\right). \hspace{1cm} (1.54)
\]

Thus we have

\[
\frac{\partial H}{\partial p} = \dot{q} \quad \frac{\partial H}{\partial q} = -\frac{\partial L}{\partial q} \quad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}
\]

\text{were } L \text{ is a function } \dot{q} \text{ and } H \text{ is a function of the corresponding } p. \text{ You should be able to show that these results (together with the Euler-Lagrange equations) yield Hamilton’s equations of motion:}

\[
\frac{dq}{dt} = \frac{\partial H(q, p, t)}{\partial p}
\]

\text{and the equation of motion are}

\[
\frac{dp}{dt} = -\frac{\partial H(q, p, t)}{\partial q}
\]

\[
\frac{dq^i}{dt} = \frac{\partial H(q, p, t)}{\partial p_i}
\]

\[
\frac{dp_i}{dt} = -\frac{\partial H(q, p, t)}{\partial q^i}
\]

- When more variables are around then we simply sum over the \(p_i\dot{q}^i\) term

\[
H(p, q, t) = \sum_i p_i \dot{q}^i(p) - L(\dot{q}(p), q, t)
\]

(1.58)
• The total derivative of the Hamiltonian satisfies
  \[ \frac{dH}{dt} = \frac{\partial H}{\partial t} \]  
  (1.61)
  so that if \( H \) is not an explicitly function of time then it is constant.

• For a (rather general) Lagrangian of the form
  \[ L = \frac{1}{2} a_{ij}(q) \dot{q}^i \dot{q}^j + b_i(q) \dot{q}^i - U(q), \]  
  (1.62)
  the momenta and velocities are related via
  \[ p_i = a_{ij} \dot{q}^j + b_i, \quad \dot{q}^i = (a^{-1})^{ij}(p_j - b_j). \]  
  (1.63)
  The Hamiltonian is
  \[ H(p, q, t) = \frac{1}{2} (a^{-1})^{ij}(p_i - b_i)(p_j - b_j) + U(q). \]  
  (1.64)
  This should be compared to the hamiltonian function in (1.46). The Hamiltonian is a function of the \( b_i \), while the hamiltonian function is not. The Hamiltonian and hamiltonian function return the same value at corresponding points where \( \dot{q} = \dot{q}(p) \), but have different functional forms.

The action principle

• The Hamiltonian can be used in the action principle to determine the equation of motion. The action takes a path in \( p, q \) space (\( p_i(t), q^i(t) \)) and returns a number
  \[ S[p(t), q(t), t] = \int dt \ (p_i \dot{q}^i - H(p, q, t)) \]  
  (1.65)
  We note \( p_i \dot{q}^i - H = L \) at corresponding points. Varying the action with \( p_i(t) \) and \( q^i(t) \) separately (keeping the ends fixed) gives the Hamiltonian equation of motion. By doing this variation you should be able to show that
  \[ \frac{dq^i}{dt} = \frac{\partial H}{\partial p_i}, \]  
  (1.66)
  \[ \frac{dp_i}{dt} = - \frac{\partial H}{\partial q^i}. \]  
  (1.67)

The Routhian

• It is often convenient to Legendre transform with respect to some of the coordinates. (This is usually convenient for the cyclic coordinates).
  Suppose we have two coordinates \( x \) and \( y \), with Lagrangian \( L(\dot{x}, x, \dot{y}, y) \). If we Legendre transform with respect to \( \dot{x} \) (replacing it with \( p_x \)), but leave \( \dot{y} \) alone:
  \[ R(p_x, x, \dot{y}, y) = p_x \dot{x}(p_x) - L(\dot{x}(p_x), x, \dot{y}, y), \]  
  (1.68)
  then \( R \) (known as the Routhian) acts like a Hamiltonian for \( (p_x, x) \), but a Lagrangian\(^2\) for \( (\dot{y}, y) \). You should be able to show that
  \[ \frac{dx}{dt} = \frac{\partial R}{\partial p_x}, \]  
  (1.69)
  \[ \frac{dp_x}{dt} = - \frac{\partial R}{\partial x} \]  
  (1.70)
  \[ \frac{d}{dt} \left( \frac{\partial R}{\partial \dot{y}} \right) = \frac{\partial R}{\partial y} \]  
  (1.71)
  Here, since the variables in \( R \) are \( p_x, x, \dot{y} \) and \( y \), the partial derivative, \( \partial R/\partial y \), means, \( (\partial R/\partial y)_{p_x} \). In the Lagrangian setup \( L(\dot{x}, x, \dot{y}, y) \), with variables \( \dot{x}, x, \dot{y} \) and \( y \), one would have \( (\partial L/\partial y)_{\dot{x}} \).

\(^2\)Technically it is actually \(-R\) that is Lagrangian for \( \dot{y}, y \), due to the fact we are subtracting \( L \) when making the Legendre transform in Eq. (1.68). Of course you could have done the following \( R = L - p_x \dot{x} \), and then it would be a Lagrangian for \( y \), but \(-R\) would be the Hamiltonian for \( x \).
The Legendre Transform as extremization in the presence of an external bias (force)

- Consider the convex function $U(x)$. Its derivative is\(^3\)

\[ dU = f_0(x) \, dx \]  

Then we define\(^4\)

\[ \hat{U}(x, f) = fx - U(x) . \]

Then the Legendre transform is the extremum (maximum or minimum) of $\hat{U}(x, f)$ for fixed $f$, i.e.

\[ V(f) = \text{extrm}_x \left( fx - U(x) \right) . \]

This means that we are to change $x$ until we reach the value $x(f)$ where $\hat{U}$ is a maximum or minimum. The value of $\hat{U}$ at this point is $V(f)$. By differentiation, the extremal point is when $f = dU/dx = f_0(x)$, which must be inverted to determine $x(f)$. Then $V(f) = fx(f) - U(x(f))$.

- We have

\[ dU = f(x) \, dx \quad \text{and} \quad dV = x(f) \, df \]

and a relation between the second derivatives

\[ \frac{d^2U}{dx^2} \frac{d^2V}{df^2} = 1 \]

- Then inverse Legendre transform returns the back the potential

\[ U(x) = \text{extrm}_f \left( fx - V(f) \right) \]

which you should prove for yourself.

- For more degrees of freedom, take $U(x_1, x_2)$ for example, the procedure works similarly. We define

\[ V(f_1, f_2) = \text{extrm}_{x_1, x_2} \left( f_1 x_1 + f_2 x_2 - U(x) \right) \]

Then

\[ dU = f_1 \, dx_1 + f_2 \, dx_2 \quad \text{and} \quad dV = x_1 \, df_1 + x_2 \, df_2 \]

Note that the matrices of second derivatives

\[ U_{ij} \equiv \frac{\partial^2U}{\partial x^i \partial x^j} \quad V^{ij} \equiv \frac{\partial^2V}{\partial f_i \partial f_j} \]

are inverses of each

\[ V^{il} U_{lj} = \delta^i_j \]

---

\(^3\)Think of $U(x)$ as the spring like potential that a particle feels. Then $f_0(x)$ is the external force that must be applied to the system so that the particle is in equilibrium at position $x$. The “internal” force that the potential gives is $f_{\text{internal}}(x) = -dU/dx$. This internal force must be counterbalanced by the applied force $f_0(x) = -f_{\text{internal}}(x)$.

\(^4\)Referring to the previous footnote $\hat{U}(x, f)$ is minus the potential in the presence of an applied external force $f$. In thermodynamics we would define the Legendre transform with $\hat{U} = U - fx$, but the overall sign leads only to minor differences. We follow the mechanics convention, $H = pv - L$, with regard to sign.
1.4 Motion in a Central Potential

Central potentials $U(r)$ and the Kepler Problem

- We have two bodies with $m_1$ and $r_1$ and $m_2$ and $r_2$, and generally take $r_1$ to be the “earth” and $r_2$ and sun. We first switch to center of mass $R$ and relative coordinates $r$

$$R = \frac{m_1 r_1 + m_2 r_2}{M}, \quad (1.82)$$

$$r = r_1 - r_2. \quad (1.83)$$

with $M = m_1 + m_2$. We have the kinetic energy

$$T = \frac{1}{2} M \dot{R}^2 + \frac{1}{2} \mu \dot{r}^2 \quad (1.84)$$

where $\mu = m_1 m_2 / (m_1 + m_2)$ is the reduced mass, and thus the Lagrangian is

$$L = \frac{1}{2} M \dot{R}^2 + \frac{1}{2} \mu \dot{r}^2 - U(|r|) \quad (1.85)$$

where $U(|r|)$ is the potential energy of the two particles.

- The overall center of mass motion does not change the orbital dynamics. We can choose $R = \dot{R} = 0$, so that the angular momentum of the center of mass is zero. Then the internal angular momentum is

$$L = \mu r \times \dot{r} \quad (1.86)$$

$L$ can be chosen to lie along the z axis so that $r$ lies in the $x, y$ plane

$$r = r (\cos \phi, \sin \phi, 0) \quad (1.87)$$

The Lagrangian neglecting the center of mass motion is

$$L = \frac{1}{2} \mu (r^2 + r^2 \dot{\phi}^2) - U(r) \quad (1.88)$$

- There are two integrals of motion for the motion in the effective potential:

$$\ell = \mu r^2 \dot{\phi}, \quad (1.89)$$

$$E = \frac{1}{2} \mu \dot{r}^2 + V_{\text{eff}}(r, \ell) \quad (1.90)$$

The effective particle with mass $\mu$ moves in the effective potential is

$$V_{\text{eff}}(r, \ell) = \frac{\ell^2}{2 \mu r^2} + U(r). \quad (1.91)$$

Given the integrals of motion $E$ and $\ell$ it is easy to determine $d\phi/d$ and $dr/dt$. From there it is straightforward to find an equation for $dr/d\phi = \dot{r}/\dot{\phi}$. Integrating $dr/d\phi$ gives the orbit for $r(\phi)$. This integral from $(r_1, \phi_1)$ to $(r, \phi)$ is

$$\phi - \phi_1 = \ell \int_{r_1}^{r} \frac{dr/r^2}{\sqrt{E - V_{\text{eff}}(r, \ell)}} \quad (1.92)$$

for an arbitrary potential $U(r)$.

- For the coulomb potential $U = -k/r$, Eq. (1.92) for $r(\phi)$ can be integrated by making the “conformal” substitution

$$u \equiv \frac{1}{r}, \quad du = \frac{dr}{r^2}, \quad (1.93)$$
1.4. MOTION IN A CENTRAL POTENTIAL

\[
\frac{1}{r} = \frac{1}{r_0} \left(1 + \epsilon \cos(\phi)\right).
\]

\( r_0 \) is known as the lattice rectum (see figure for geometric meaning), and \( \epsilon \) is known as the eccentricity of the ellipse, which is a measure of how much the orbit deviates from a circle. A convenient summary of the elliptic geometry is given in Fig. 1.1. The parameters of the ellipse \( r_0 \) and \( \epsilon \) are determined by the integrals of motion, \( E \) and \( \ell \). The lattice rectum is determined by the angular momentum, \( r_0 = \ell^2/\mu k \). The eccentricity \( \epsilon \) is determined by the excitation energy above the minimum of \( V_{\text{eff}} \) (with fixed \( \ell \)). More explicitly \( \epsilon = \sqrt{1 + \frac{E}{\epsilon_0}} \), with \( \epsilon_0 = \frac{\ell^2}{2\mu r_0^2} \). When the energy of the orbit is at its minimum, \( E = V_{\text{min}} = -\epsilon_0 \), then the eccentricity is zero and the radius is constant, i.e. the orbit is circular.

- The Coulomb potential has a characteristic scale \( r_0 \sim \ell^2/\mu k \) when the potential \( k/r_0 \) and kinetic \( \ell^2/\mu r_0^2 \) are the same order of magnitude. Indeed, for a circular orbit of radius \( r_0 \), one shows by freshman physics that the radius is determined by the angular momentum, \( r_0 = \ell^2/\mu k \). For such a circular orbits the kinetic energy is \( \epsilon_0 = \ell^2/2\mu r_0^2 \) and is minus-half the potential \( U = -k/r_0 = -2\epsilon_0 \). The total energy (kinetic+potential) is \( E = -\epsilon_0 \) where

\[
\epsilon_0 = \frac{\ell^2}{2\mu r_0^2} = \frac{k}{2r_0},
\]

which explains the notation for the parameters in the previous item.

- For the Newton potential \( U = -k/r \) and the spherical harmonic oscillator \( U = \frac{1}{2}kr^2 \) the orbits are closed (Bertrand’s theorem). For no other central potentials are the orbits closed. The closed orbits are a consequence of an additional symmetry which we will discuss later.

**Cross sections and scattering**

- When considering the scattering problem we are interested in computing the scattering angle \( \theta \) (the angle of deflection) for given energy \( E \) and impact parameter \( b \). Here the impact parameter \( b \) is the transverse distance at large \( r \) from the target and is another way to record the angular momentum. At larger \( r \) the velocity is constant, \( E = \frac{1}{2}mv^2 \), and the angular momentum is

\[
\ell = mvr \sin \theta = mveb = \sqrt{2mE}b
\]
• The scattering angle $\theta(b)$ is shown below:

A particle comes in with impact parameters $b$ (or angular momentum $\ell$) and energy $E$, and is deflected by angle $\theta(b,E)$. From our mechanical perspective we find it easiest to compute the change in the angle $\phi$ as the particle propagates from its distance of closest approach $r_{\text{min}}$ up to infinity. This is (the second) angle $\psi$ in the figure above. It is related to $\theta(b,E)$ by simple geometry.

$$\theta(b) = \pi - 2\psi. \quad (1.97)$$

We have from Eq. (1.92)

$$\Delta \phi = \psi = \frac{\ell}{\sqrt{2m}} \int_{r_{\text{min}}}^{\infty} \frac{dr}{r^2} \frac{dV}{E - V_{\text{eff}}(r)}^{1/2}. \quad (1.98)$$

For the Coulomb problem $U = k/r$ this integration is straightforward with the substitution $u = 1/r$, and yields $\tan(\psi)$ and since $\psi = \pi/2 - \theta/2$

$$\cot(\theta/2) = \frac{2Eb}{k}. \quad (1.99)$$

• The scattering problem is usually phrased in terms of cross section:

(i) Consider a beam of particles of luminosity $\mathcal{L}$. $\mathcal{L}$ is the number of particles crossing the target per area per time, and is also called the incident flux or intensity.

(ii) The number of incoming particles which scatter per time $d\Gamma$ with impact parameter between $b$ and $db$ is $d\Gamma = \mathcal{L}2\pi b|db|$. We put absolute values because we think of $db$ as an positive interval.

(iii) The number of incoming particles per time (or rate $d\Gamma$) which then end up at in ring of solid angle $d\Omega = 2\pi \sin(\theta) d\theta$ per time is

$$d\Gamma = \mathcal{L} \frac{b}{\sin \theta} \frac{|db|}{|d\theta|} d\Omega. \quad (1.100)$$

So the scattering rate per solid angle is

$$\frac{d\Gamma}{d\Omega} = \mathcal{L} \frac{b}{\sin \theta} |db| |d\theta|. \quad (1.101)$$

The cross section is by definition the scattering rate divided by the incident flux

$$\frac{d\sigma}{d\Omega} = \frac{1}{\mathcal{L}} \frac{d\Gamma}{d\Omega} = \frac{b}{\sin \theta} \left| \frac{db}{d\theta} \right|. \quad (1.102)$$

(iv) The cross section has units of area and gives a measure of the effective size of the target. It is usually measured in barns, $1$ barn $= 10^{-24}$ cm$^2$. 

• The scattering problem is usually phrased in terms of cross section:
For the Coulomb problem, we can different $d\theta/db$ (Eq. (1.99)) and use it in Eq. (1.102) to determine the Rutherford cross section

$$
\frac{d\sigma}{d\Omega} = \left( \frac{k}{4E} \right)^2 \frac{1}{\sin^4(\theta/2)} \sim \frac{1}{\theta^4},
$$

(1.103)

which is inversely proportional to $1/\theta^4$ at small angles.
1.5 Constraints

Lagrange multipliers

First we considered minimizing $U(x, y)$ subject to a constraint $Q(x, y) = 0$. We said that we should instead minimize

$$
\hat{U}(x, y, \lambda) = U(x, y) - \lambda Q(x, y).
$$

$\lambda$ is known as a Lagrange multiplier\(^5\). This leads to the conditions

$$
d\hat{U}(x, y) = \left( \frac{\partial U}{\partial x} - \lambda \frac{\partial Q}{\partial x} \right) dx + \left( \frac{\partial U}{\partial y} - \lambda \frac{\partial Q}{\partial y} \right) dy - Q d\lambda = 0
$$

where the terms in front of $dx$, $dy$, and $d\lambda$ should be set to zero. We explained that $Q$ can be thought of as a generalized coordinate, and $\lambda$ is a generalized force conjugate to $Q$. This is just like adding an external force. For instance if I have a potential $U(x, y)$ and add an external force $f$ in the $x$ direction then the new potential is

$$
\hat{U}(x, y, f) = U(x, y) - fx.
$$

The forces of constraint in the $x$ and $y$ directions are

$$
F_x = \lambda \partial_x Q, \quad F_y = \lambda \partial_y Q.
$$

The setup easily generalizes to more coordinates and more constraints. For coordinates $x^A$ and constraints $Q^\alpha(x^A)$ with $\alpha = 1 \ldots m$, if we want to minimize $U(x^A)$ subject to these constraints, we instead extremize

$$
\hat{U}(x^A) = U(x^A) - \lambda_\alpha Q^\alpha(x^A)
$$

requiring that $d\hat{U} = 0$, i.e. require

$$
\frac{\partial \hat{U}}{\partial x^A} = 0 \quad \frac{\partial \hat{U}}{\partial \lambda_\alpha} = 0
$$

The forces of constraint in the $x^A$ direction are

$$
F_A = \lambda_\alpha \frac{\partial Q^\alpha}{\partial x^A}
$$

Newton’s Laws and Lagrange with constraints

Consider Newton’s Laws for particles with positions $r_a$. For simplicity consider just one constraint.

$$
Q(r_a) = 0
$$

Then

$$
dQ = \nabla r_a Q \cdot dr_a = 0
$$

The forces of constraints $F_a^C$ do no work

$$
F_a^C \cdot dr^a = 0
$$

\(^5\)The sign in front of $\lambda$ is irrelevant. The choice here is so that $\lambda$ corresponds to the generalized force in the direction of increasing $Q$, compare to Eq. (1.106). When we consider constraints in the Lagrangian, $L = T - U$, the multipliers will then come with a plus sign $L = T - U + \lambda Q$.\}
Thus, we make take $F_C^a$ to be proportional to the gradient of $Q$

$$F_C^a = \lambda \nabla_{r_a} Q$$  \hspace{1cm} (1.116)

Then Newton’s Laws read

$$\frac{dp_a}{dt} = F_a^{\text{ext}} + \lambda \nabla_{r_a} Q.$$  \hspace{1cm} (1.117)

Then Newton’s Law ($F = ma$) and the constraint, determine the accelerations of the particles and the magnitude of the forces of constraint, i.e. $\lambda$.

- You should do some simple problems on Attwood’s machines (see below) to convince yourself that we are always solving Eq. (1.117) when doing Freshmann physics problems.

- In the Lagrangian formalism we add some lagrange multipliers to enforce the constraints. Instead of extremizing $L(\dot{r}_a, r_a)$, one extremizes $\hat{L}(\dot{r}_a, r_a, \lambda) = L + \lambda Q$, where $\lambda$ is like an extra coordinate. The Euler-Lagrange equations for $\hat{L}$ are

$$\frac{d}{dt} \left( \frac{\partial \hat{L}}{\partial \dot{r}_a} \right) = \frac{\partial \hat{L}}{\partial r_a}$$  \hspace{1cm} (1.120)

$$0 = Q$$  \hspace{1cm} (1.121)

- If there are more constraints $Q^\alpha$, simply make the replacement $\lambda Q \rightarrow \lambda_\alpha Q^\alpha$ in the lagrangian formalism. In the Newtonian formalism the force of constraint on the $a$-th particle is

$$F_a = \lambda_\alpha \nabla_{r_a} Q^\alpha.$$  \hspace{1cm} (1.122)

- **Attwood machine.** Consider two masses $m_1$ and $m_2$ hanging over a massless pulley (you know the problem!). We have two coordinates $z_1$ and $z_2$ where $z_1$ and $z_2$ are the distances below the pulley (increasing $z$ means further down). The constraint is

$$Q = z_1 + z_2 - L$$  \hspace{1cm} (1.123)

The hatted Lagrangian is

$$\hat{L} = \frac{1}{2} m_1 \dot{z}_1^2 + \frac{1}{2} m_2 \dot{z}_2^2 + m_1 g z_1 + m_2 g g z_2 + \lambda (z_1 + z_2 - L)$$  \hspace{1cm} (1.124)

Newton’s or Lagranges’ equation of motion are

$$m_1 a_1 = m_1 g + \lambda$$  \hspace{1cm} (1.125)

$$m_2 a_2 = m_2 g + \lambda$$  \hspace{1cm} (1.126)

$$z_1 + z_2 = L$$  \hspace{1cm} (1.127)

Which are easily solved for $a_1$, $a_2$ and $\lambda$, using that Eq. (1.127) implies by differentiating that $a_1 + a_2 = 0$. Solving these equations gives $\lambda$ negative, i.e. the force is up not down. The case when the pulley has mass in the Lagrangian formalism is suggested as an exercise.

---

6Perhaps we should write it a bit more explicitly. The coordinates of $r_a$ are $r^i_a$ with $i = x, y, z$. We mean

$$\frac{d}{dt} \left( \frac{\partial \hat{L}}{\partial \dot{r}_a^i} \right) = \frac{\partial \hat{L}}{\partial r_a^i}$$  \hspace{1cm} (1.118)

$$\frac{d}{dt} \left( \frac{\partial \hat{L}}{\partial \lambda} \right) = \frac{\partial \hat{L}}{\partial \lambda}$$  \hspace{1cm} (1.119)

The equation $0 = Q$ follows from the equation for $\lambda$, which simply enforces the constraint.