Physics 235 Chapter Notes

Frank Wolfs

Abstract

These lecture notes contain the material I cover when teaching PHY 235 at the University of Rochester. The text book used in this course is *Classical Dynamics of Particles and Systems*, *Thornton and Marion (Fifth Edition)* The figures that appear in these lecture notes are figures that appear in the text book, unless otherwise noted.

This document will be updated as additional Chapters are converted to Latex.

The URL of the course website is: http://teacher.pas.rochester.edu/PHY235/Phy235HomePage.shtml.

Contents

1 Matrices, Vectors, and Vector Calculus

In this Chapter, we will focus on the mathematical tools required for the course. The main concepts that will be covered are:

1. Coordinate transformations
2. Matrix operations
3. Scalars and vectors
4. Vector calculus
5. Differentiation and integration

1.1 Coordinate transformations

In order to be able to specify the position of a point P we first must specify the coordinate system that will be used. The coordinates of point P will be a function of the coordinate system being used, and coordinate transformations allow us to define the relation between the coordinates of point P in different coordinate systems.

Different types of coordinate systems are used for different applications. The most commonly used coordinate systems are:

1. **Cartesian coordinate systems**. These systems consist out of three perpendicular coordinate axes, called the x, y, and z (or x_1, x_2, and x_3) axes. The coordinates of a point P are usually specified by three coordinates (x, y, z) or (x_1, x_2, x_3).

   Note: the coordinate axes in a Cartesian coordinate system are usually independent of time.

   Note 2: the direction of the third coordinate axis is defined by the directions of the first two coordinate axes and the right-hand rule.

2. **Spherical coordinate systems**. Spherical coordinate systems are most often used when the system under consideration has spherical symmetry. The origin of the coordinate system is chosen to coincide with the point of spherical symmetry. The position of a point P is determined by specifying the distance r from the origin of the coordinate system and the polar and azimuthal angles θ and ϕ.

   Note: we still need to define a Cartesian coordinate system to define the origin and the two angles.

   Note 2: the unit vectors associated with the position vector and the angles will be a function of time if the position described is time dependent.

3. **Cylindrical coordinate systems**. Cylindrical coordinate systems are most often used when the system under consideration has cylindrical symmetry. The coordinate system is characterized by the axis of cylindrical symmetry, which is usually called the z axis. The position of a point P is determined by specifying the distance r to the z axis, the azimuthal angle ϕ, and the z coordinate.

   Note: we still need to define a Cartesian coordinate system to define the origin and the azimuthal angle.

   Note 2: the unit vectors associated with the position vector and the azimuthal angle will be a function of time if the position described is time dependent.

A coordinate system is not uniquely defined. For example, the coordinates of a point P in a Cartesian coordinate system depend on the choice of coordinate axes. Different choices will result in different coordinates. **Coordinate transformations** are used to transform the coordinates between coordinate systems. There are a number of different types of coordinate transformations we will encounter in this course: translation, rotation, and the standard Lorentz transformation. In this Chapter, we will only focus on the rotational transformation.

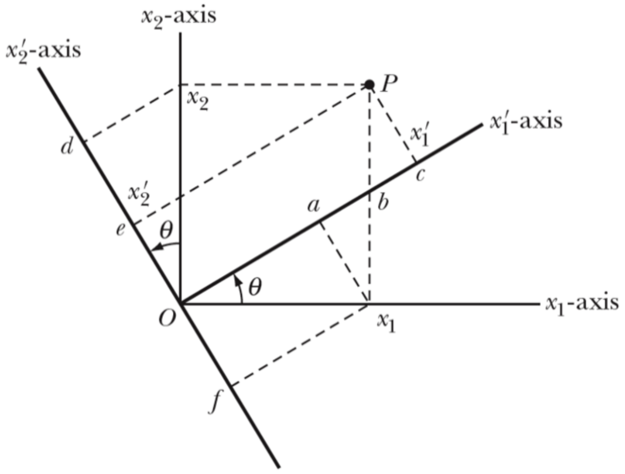


Figure 1.1: Two different coordinate systems used to represent the position of *P*.

Consider two coordinate systems, related to each other via a transformation around the x_3 axis, as shown in Fig. 1.1. The relation between the coordinates of *P* in the two coordinate systems can be written as

$$x_i{}' = \sum_{j=1}^{3} \lambda_{ij} x_j \tag{1.1}$$

where

$$\lambda_{ij} = \cos\left(x_i{}', x_j\right) \tag{1.2}$$

is the cosine of the angle between the $x_i{}'$-axis and the x_j-axis (also called the **direction cosine**). The coordinate transformation is most often written in matrix notation:

$$\vec{X}' = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{pmatrix} \vec{X} \tag{1.3}$$

Consider the two coordinate systems shown in Figure 1. Since the two coordinate systems are related to each other via a rotation around the x_3-axis we conclude that:

1. The direction cosine between the $x_3{}'$-axis and the x_1- and x_2-axes will be zero since the angle between the $x_3{}'$-axis and the x_1- and x_2-axes is 90°). Thus: $\lambda_{31} = \lambda_{32} = 0$.
2. The direction cosine between the $x_3{}'$-axis and the x_3-axis will be one since the angle between the $x_3{}'$-axis and the x_3-axis is 0°). Thus $\lambda_{33} = 1$.
3. The direction cosine between the $x_1{}'$- and $x_2{}'$-axes and the x_3-axis will be zero (angle is 90°). Thus: $\lambda_{13} = \lambda_{23} = 0$.
4. The direction cosines between the $x_1{}'$- and the x_2-axis is $\cos(\pi/2-\theta)$ and the direction cosine between the $x_2{}'$- and the x_1-axis is $\cos(\pi/2+\theta)$. Thus: $\lambda_{12} = \cos(\pi/2-\theta) = \sin(\theta)$ and $\lambda_{21} = \cos(\pi/2+\theta) = -\sin(\theta)$.
5. The direction cosines between the $x_1{}'$- and the x_1-axes and between the $x_2{}'$- and the x_2-axes is $\cos(\theta)$. Thus: $\lambda_{11} = \lambda_{22} = \cos(\theta)$.

The coordinate transformation for this particular transformation is thus given by

$$\vec{X}' = \begin{pmatrix} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \vec{X} \tag{1.4}$$

Although each rotation matrix has 9 parameters, only 3 are truly independent. This can be seen by realizing that in order to specify a rotation we need to specify the rotation axis (which requires the specification of a polar and azimuthal angle) and the rotation angle. Consider some of the relations we can determine between the parameters of the rotation matrix:

1. The rotation preserves the length of a vector. Rotating a unit vector will produce another unit vector. This requires that

$$\sqrt{{\lambda_{11}}^2 + {\lambda_{21}}^2 + {\lambda_{31}}^2} = 1 \tag{1.5}$$

$$\sqrt{{\lambda_{12}}^2 + {\lambda_{22}}^2 + {\lambda_{32}}^2} = 1 \tag{1.6}$$

$$\sqrt{{\lambda_{13}}^2 + {\lambda_{23}}^2 + {\lambda_{33}}^2} = 1 \tag{1.7}$$

2. The rotation preserves the angle between two vectors. Since the angle between the coordinate axes is 90°, the angle between these axes after transformation must also be 90°. This requires that

$$\lambda_{11}\lambda_{12} + \lambda_{21}\lambda_{22} + \lambda_{31}\lambda_{32} = 0 \tag{1.8}$$

$$\lambda_{11}\lambda_{13} + \lambda_{21}\lambda_{23} + \lambda_{31}\lambda_{33} = 0 \tag{1.9}$$

$$\lambda_{12}\lambda_{13} + \lambda_{22}\lambda_{23} + \lambda_{32}\lambda_{33} = 0 \tag{1.10}$$

These six equations can be combined in the following manner

$$\sum_{j=1}^{3} \lambda_{ij}\lambda_{kj} = \delta_{ik} \tag{1.11}$$

This equation is called the **orthogonal condition**, and is only satisfied if the coordinate axes are mutually perpendicular.

1.2 Matrix operations

When we use the rotation matrix to carry out coordinate transformations, we carry out a matrix operation. There are several important facts to remember about matrix operations:

1. The **unit matrix** is defined as a matrix for which the diagonal values are 1 and the non-diagonal values are 0. When the unit vector operates on a vector, the result is the same vector.
2. The **inverse of a matrix** is defined such that when it operates on the original matrix, the result is the unit matrix.
3. A **transposed matrix** is derived from the original matrix by interchanging the rows and columns.
4. When we combine coordinate transformations, we can obtain the resulting transformation by multiplying the rotation matrices.

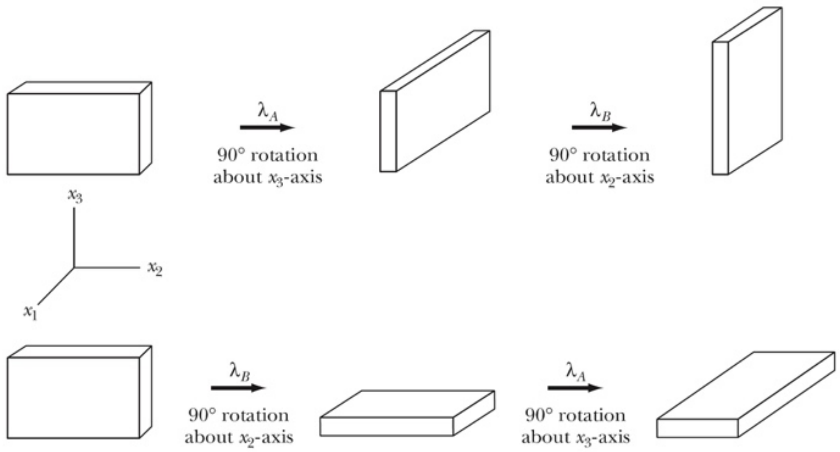


Figure 1.2: The order of transformation matters.

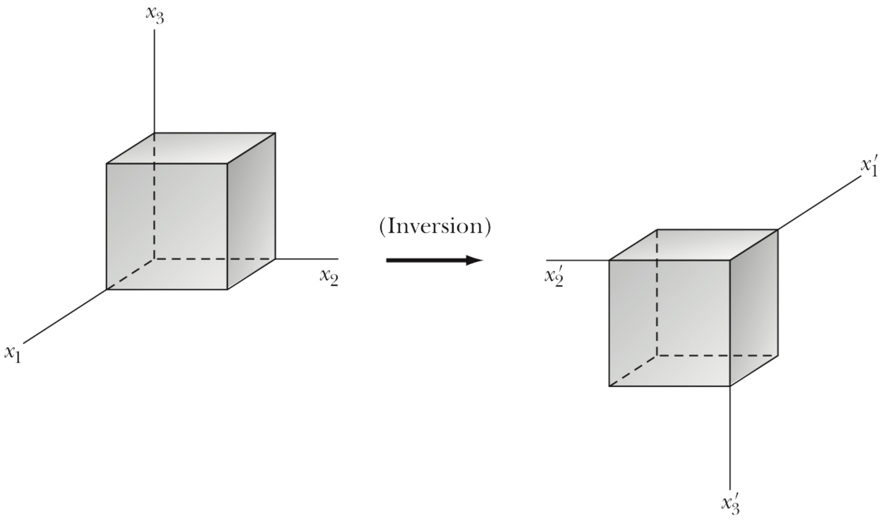


Figure 1.3: An example of an inversion.

5. When we multiply rotation matrices we must realize that **the order of the multiplication matters**. See for example Fig. 1.2.

6. A matrix inversion is a transformation that results in the reflection through the origin of all axes. We can not find any series of rotations that result in an inversion. Matrix inversions are examples of the so-called **improper rotations**, and are characterized by a matrix with a determinant equal to -1. **Proper rotations** are those rotations that are characterized by a matrix with a determinant equal to +1. An example of an inversion is shown in Fig. 1.3. Note that the coordinate axes of the coordinate frame after the inversion are now consistent with the right-hand rule.

1.3 Vectors and Scalars

Consider the following coordinate transformation

$$x_i{}' = \sum_{j=1}^{3} \lambda_{ij} x_j \tag{1.12}$$

where

$$\sum_{j=1}^{3} \lambda_{ij}\lambda_{kj} = \delta_{ik} \tag{1.13}$$

Vectors and scalars are defined based on what happens as a result of such coordinate transformations:

1. If a quantity is unaffected by this transformation, it is called a **scalar**.
2. If a set of three quantities transforms in the same manner as the coordinates of a point P, these quantities are the components of what we call a **vector**.

Note:

1. These definitions define scalars and vectors in a very different way from what you may have been used to. For example, the current definition of a vector makes no reference to the geometrical interpretation the vector.
2. You will encounter parameters in Physics that look like vectors, but do not transform like vectors under certain operations (a pseudo vector is an example). Just having a magnitude and a direction is not sufficient to define a vector!

Vector and scalar addition and multiplication have a number of properties in common:

1. Both satisfy the commutative law: the order of addition does not change the final result.
2. Both satisfy the associative law: when we determine the sum of more than two vectors or scalars, the final results will not depend on which pair of vectors or scalars we add first.
3. The product of a scalar with a scalar transforms like a scalar (**and is thus a scalar**).
4. The product of a vector with a scalar transforms like a vector (**and is thus a vector**).

The following operations are unique to vectors and do not have equivalents for scalars:

1. **The Scalar product**.

   The scalar product between two vectors $\vec{A}$ and $\vec{B}$ is defined as

   $$\vec{A} \cdot \vec{B} = \sum_{i=1}^{3} A_i B_i \tag{1.14}$$

   It can be shown that the scalar product as defined above, is equal to the product of the magnitudes of the two vectors and the direction cosine between them:

   $$\vec{A} \cdot \vec{B} = \sum_{i=1}^{3} A_i B_i = AB\cos\left(\vec{A},\vec{B}\right) \tag{1.15}$$

   In order to show that the scalar product behaves like a scalar, we must show that the scalar product between $\vec{A}$ and $\vec{B}$ is the same as the scalar product between $\vec{A}'$ and $\vec{B}'$.

   The scalar product also satisfies the **commutative** and the **distributive** laws:

   $$\vec{A} \cdot \vec{B} = \vec{B} \cdot \vec{A} \tag{1.16}$$

   $$\vec{A} \cdot \left(\vec{B} + \vec{C}\right) = \vec{A} \cdot \vec{B} + \vec{A} \cdot \vec{C} \tag{1.17}$$

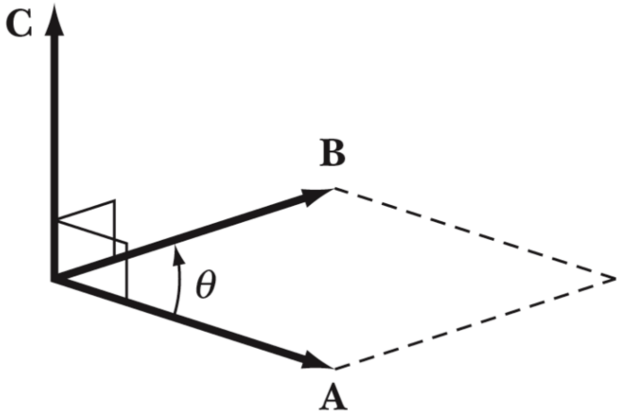


Figure 1.4: Properties of the vector product between the vectors $\vec{A}$ and $\vec{B}$.

2. **The Vector product**.

   The vector product between two vectors $\vec{A}$ and $\vec{B}$ is a third vector $\vec{C}$, defined as

   $$\vec{C} = \vec{A} \times \vec{B} = \begin{vmatrix} \hat{x}_1 & \hat{x}_2 & \hat{x}_2 \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \tag{1.18}$$

   In order to show that the vector product behaves like a vector, we must show that the vector product transforms like a vector under a coordinate transformation.

   The geometrical interpretation of the vector product is shown in Fig. 1.4. The magnitude of the vector product is the area of the parallelogram defined by the vectors $\vec{A}$ and $\vec{B}$ and it is directed in a direction perpendicular to the plane defined by the vectors $\vec{A}$ and $\vec{B}$ (the right-hand rule defines the direction).

   There are many properties of the vector product. Some of them are listed here (see the text book for a more complete listing):

   $$\vec{A} \times \vec{B} = -\vec{B} \times \vec{A} \tag{1.19}$$

   $$\vec{A} \times \left(\vec{B} \times \vec{C}\right) = \left(\vec{A} \cdot \vec{C}\right)\vec{B} - \left(\vec{A} \cdot \vec{B}\right)\vec{C} \tag{1.20}$$

   $$\vec{A} \cdot \left(\vec{B} \times \vec{C}\right) = \vec{B} \cdot \left(\vec{C} \times \vec{A}\right) = \vec{C} \cdot \left(\vec{A} \times \vec{B}\right) \tag{1.21}$$

1.4 Differentiation and Integration

Two important operations on both scalars and vectors are differentiation and integration. These operations are used to define important mechanical quantities (such as velocity and acceleration), and a thorough understanding of operations involving differentiation and integration is required in order to succeed in this course.

1. **Scalar Differentiation**.

   We can differentiate vectors and scalars with respect to a scalar variable s.

   (a) The result of the differentiation of a scalar with respect to another scalar variable will be another scalar. The result of the differentiation will be independent of the coordinate system.

(b) The result of the differentiation of a vector function with respect to a scalar variable will be another vector. The resulting vector will be directed tangential to the curve that represents the function.

2. **Scalar Differentiation in different coordinate systems**.

An important scalar variable used in differentiations is the time t. Based on the position vector, we can obtain the velocity and acceleration vectors by differentiating the position vector once and twice, respectively, with respect to time. If Cartesian coordinates are being used, the axes are independent of time, and differentiation the position vector with respect to time is equivalent to differentiating the individual components with respect to time:

$$\vec{v} = \dot{\vec{r}} = \sum_i \frac{dx_i}{dt}\hat{x}_i \tag{1.22}$$

$$\vec{a} = \dot{\vec{v}} = \ddot{\vec{r}} = \sum_i \frac{d^2x_i}{dt^2}\hat{x}_i \tag{1.23}$$

The situation is more complicated if we are using spherical or cylindrical coordinates. Consider for example the motion of an object shown in Fig. 1.5. When the object moves during a time dt from $P(1)$ to $P(2)$, the spherical unit vectors change too, as shown in Fig. 1.5:

$$d\hat{r} = \hat{r}(t+dt) - \hat{r}(t) \tag{1.24}$$

$$d\widehat{\theta} = \widehat{\theta}(t+dt) - \widehat{\theta}(t) \tag{1.25}$$

Based on the definition of the unit vectors in the spherical coordinate systems we can conclude:

$$d\hat{r} = (d\theta)\widehat{\theta} \tag{1.26}$$

$$d\widehat{\theta} = -(d\theta)\hat{r} \tag{1.27}$$

By dividing each side by dt we obtain the following relations:

$$\dot{\hat{r}} = \frac{d\hat{r}}{dt} = \left(\frac{d\theta}{dt}\right)\widehat{\theta} = \dot{\theta}\widehat{\theta} \tag{1.28}$$

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

Using these relations we can calculate the velocity and acceleration:

$$\vec{v} = \frac{d}{dt}(r\hat{r}) = \frac{dr}{dt}\hat{r} + r\frac{d\theta}{dt}\widehat{\theta} = \dot{r}\hat{r} + r\dot{\theta}\widehat{\theta} \tag{1.30}$$

$$\vec{a} = \frac{d}{dt}\left(\dot{r}\hat{r} + r\dot{\theta}\widehat{\theta}\right) = \left(\ddot{r} - r\dot{\theta}^2\right)\hat{r} + \left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right)\widehat{\theta} \tag{1.31}$$

Other relations for spherical and cylindrical coordinates can be found in the textbook.

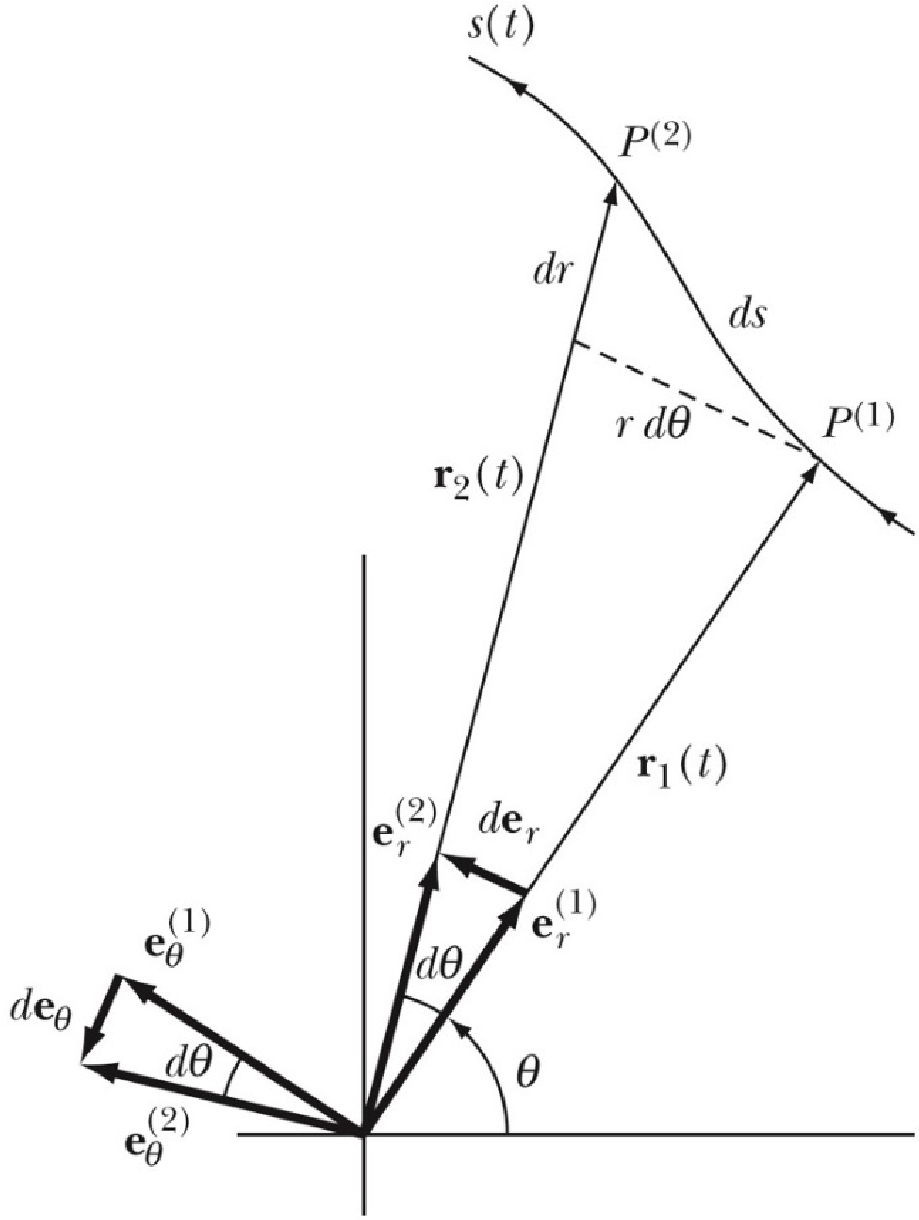


Figure 1.5: Motion of an object described in terms of spherical coordinates.

3. **Vector Differential Operator**.

   A very important operator in this course will be the gradient operator. It operates on a scalar function and the result of the operation is a vector. Consider a scalar function ϕ that is a function of the Cartesian coordinates. The value of the scalar function at a point P in two different coordinate systems must be the same:

$$\phi'(x_1{}', x_2{}', x_3{}') = \phi(x_1, x_2, x_3) \tag{1.32}$$

   The coordinates in the two different coordinate systems are connected to each other via a rotation matrix:

$$x_i{}' = \sum_{j=1}^{3} \lambda_{ij} x_j \tag{1.33}$$

   When we differentiate the scalar function we find the following relation:

$$\frac{\partial \phi'}{\partial x_i{}'} = \sum_{j=1}^{3} \lambda_{ij} \frac{\partial \phi}{\partial x_j} \tag{1.34}$$

   As we can see, the components of the differential of the scalar function transform like a vector, and the components can thus be considered the components of a vector we call the **gradient** of a scalar function:

$$grad = \vec{\nabla} = \sum_{j=1}^{3} \hat{x}_j \frac{\partial}{\partial x_j} \tag{1.35}$$

   Other important operators are defined in terms of the gradient operator:

$$\text{grad}\phi = \vec{\nabla}\phi = \sum_{j=1}^{3} \frac{\partial \phi}{\partial x_j} \hat{x}_j \tag{1.36}$$

$$\text{div}\vec{A} = \vec{\nabla} \cdot \vec{A} = \sum_{j=1}^{3} \frac{\partial A_j}{\partial x_j} \tag{1.37}$$

$$\text{curl}\vec{A} = \vec{\nabla} \times \vec{A} = \sum_{i,j,k} \varepsilon_{ijk} \frac{\partial A_k}{\partial x_j} \hat{x} \tag{1.38}$$

$$\vec{\nabla}^2 \phi = \sum_{j=1}^{3} \frac{\partial^2 \phi}{\partial x_j{}^2} \tag{1.39}$$

   The operator shows in Eq. 1.39 is called the **Laplacian**.

   The gradient of a scalar function has the following properties:

   (a) The gradient of a scalar function at point P is directed normal to the lines or surfaces for which the scalar function is constant.

   (b) The gradient of a scalar function at point P is directed in the direction of maximum change in the scalar function.

4. **Integration**.

   The opposite of differentiation is integration. Both scalar and vector functions can be integrated, and we can encounter volume, surface, and line integration:

   (a) **Volume integration of a vector**. When we integrate a vector over a volume, the result is another vector with components obtained by volume integration of the components of the original vector.

   $$\int_V \vec{A} dv = \begin{pmatrix} \int_V A_1 dv \\ \int_V A_2 dv \\ \int_V A_3 dv \end{pmatrix} \tag{1.40}$$

   (b) **Surface integration of a vector**. The surface integral of a vector function is given by the integral of its component perpendicular to the surface. When we integrate a vector over a surface, the result is a scalar.

   $$\int_S \vec{A} \cdot d\vec{a} = \int_S \left(\vec{A}\hat{n}\right) da \tag{1.41}$$

   (c) **Line integration of a vector.** The line integration of a vector is given by the integral of the component of the vector along the path (does not need to be a straight line).

   $$\int_{Line} \vec{A} \cdot d\vec{s} = \int_{Line} \sum_i A_i dx_i \tag{1.42}$$

   Various theorems relate volume, surface, line integrals of vectors. Some of the most important theorems are:

   (a) **Gauss's theorem for volume integrals**:

   $$\int_S \vec{A} \cdot d\vec{a} = \int_V \left(\vec{\nabla}\vec{A}\right) dv \tag{1.43}$$

   Note: the surface integral of a vector function is replaced by the volume integral of a scalar function.

   (b) **Stoke's theorem for line integrals**:

   $$\int_{Line} \vec{A} \cdot d\vec{s} = \int_S \left(\vec{\nabla} \times \vec{A}\right) d\vec{a} \tag{1.44}$$

   Stoke's theorem is most useful if it reducing a two-dimensional surface integral to a one-dimensional line integral.

2 Newtonian Mechanics – Single Particle

In this Chapter we will review what Newton's laws of mechanics tell us about the motion of a single particle. Newton's laws are only valid in suitable reference frames, and we will discuss what makes a reference frame a suitable reference frame. We will also review the various conservation laws you should have already encountered in your introductory physics course. In addition, we will discuss the limits of Newtonian mechanics.

2.1 Newton's Laws

The following three laws of Newtonian mechanics should have been discussed in your introductory physics course:

1. **Newton's First Law: A body remains at rest or in uniform motion unless acted upon by a force.** This law does not tell us very much about the concept of force, except what we mean with zero force: if we see a body at rest or in uniform motion, we know that the net force acting on the body is zero.
2. **Newton's Second Law: A body acted upon by a force moves in such a manner that the time rate of change of its linear momentum equals the force.** Newton defined the linear momentum of a particle of mass m moving with a velocity v as mv. The second law can thus be used to define the force $\vec{F} = d\,(m\vec{v})/dt$. This definition is of course only useful if the mass and the velocity of a particle are defined.
3. **Newton's Third Law: If two bodies exert forces on each other, these forces are equal in magnitude and opposite in direction.** This law is only true if the force acting between the bodies is directed along the line connecting the bodies (these forces are called central forces). Forces that are velocity dependent are in general non-central forces and do not satisfy Newton's Third Law. Conservation of linear momentum is a direct consequence of the third law. If $\vec{F}_1 = -\vec{F}_2$ then $d(m_1v_1)/dt$ = - $d(m_2v_2)/dt$. This equation can be rewritten as $d(m_1v_1 + m_2v_2)/dt = 0$ or $m_1v_1 + m_2v_2 =$ constant.

2.2 Reference Systems

Newton's laws are only valid in an appropriate reference frame. We can use this requirement to define inertial reference systems: **an inertial reference frame is a reference frame in which Newton's laws are valid**. If Newton's laws are valid in one reference frame, they are also valid in any reference frame in uniform motion with respect to the first frame. In order to be able to describe a free particle (a particle on which no force is acting) in a reference frame, the reference frame must satisfy the following conditions:

1. The equation of motion of the particle should be independent of the position of the origin of the coordinate system.
2. The equation of motion of the particle should be independent of the orientation of the coordinate system.
3. Time must be homogeneous (the velocity of a free particle must be constant).

2.3 Single-Particle Motion

If we know the force acting on the particle, we can use Newton's second law to describe its motion:

$$\frac{d\vec{v}}{dt} = \frac{1}{m}\vec{F} \tag{2.1}$$

Note: $\vec{F}$ must be the total force acting on the particle.
As long as the force is constant, we can usually obtain an analytical expression for the velocity and/or the trajectory of the particle. The situation becomes more complicated when the force is time dependent,

velocity dependent, and/or position dependent. In those circumstances we may need to rely on numerical methods to predict the motion of the particle. The second law can be rewritten as

$$d\vec{v} = \vec{v}(t+dt) - \vec{v}(t) = \frac{dt}{m}\vec{F} \tag{2.2}$$

The velocity at time $t + dt$ can thus be determined from the velocity at time t using the following relation:

$$\vec{v}(t+dt) = \vec{v}(t) + \frac{dt}{m}\vec{F} \tag{2.3}$$

This relation can be used to determine the velocity as function of time, if we know 1) the velocity at one specific time and 2) the force $\vec{F}$ is known (**but does not need to be constant**). Once we know the velocity as function of time, we can determine the position as function of time:

$$\vec{r}(t+dt) = \vec{r}(t) + \vec{v}(t)\,dt \tag{2.4}$$

Let us illustrate the use of numerical methods by focusing on projectile motion. If the gravitational force is the only force present, we can express the motion of the particle analytically. By comparing the results of a numerical calculation with the analytical solution, we can study the limitations of the numerical approach. Since the gravitational force is acting in the vertical direction (along the y axis), the force will have no effect on the motion in the horizontal direction (along the x axis):

$$v_x(t+dt) = v_x(t) \tag{2.5}$$

$$v_y(t+dt) = v_y(t) - g\,dt \tag{2.6}$$

$$x(t+dt) = x(t) + v_x(t)\,dt \tag{2.7}$$

$$y(t+dt) = y(t) + v_y(t)\,dt \tag{2.8}$$

Many different programs can be used to study the evolution of these equations. No matter what approach is being used, the most critical choice the user will have to make is the size of the step size dt. In the case of a constant force, the expressions for the velocity at time $t + dt$ are correct, independent of the step size dt. However, the expressions for the position at time $t + dt$ are only correct if the velocity is constant over the period between time t and time $t + dt$. This is a reasonable approximations if the step size dt is small, but for a large step size dt this is clearly a poor approximation (and large errors will result). A very small step size will increase the computing time and may lead to rounding errors.

On the Physics 235 website you can find an example of a study of projectile motion using numerical solutions (see http://teacher.pas.rochester.edu/PHY235/ComputingTools/ComputingToolsIndex.htm). Using GlowScript (Phy235-ProjectileMotion), we can study the effect of the choice of the step size dt. Consider the case of projectile motion, starting at time t = 0 s at the origin of our coordinate system with a velocity of 600 m/s and a launch angle of 60°. Figure 2.1(left) shows a comparison between the results of the analytical calculation of the trajectory of the projectile (red data points) and the results of the numerical calculation of the trajectory (blue data points) with a time step of 10 s. The difference between these two calculations, defined as in $y_{analytical} - y_{numerical}$, is shown in Fig. 2.1(right). There is clearly a significant difference between the numerical and analytical calculations. Figure 2.2 shows the results of the same calculations as shown in Fig. 2.1, except that the time step was changed to 1 s. As a result, there are clearly more data points on the trajectory. However, the more important difference is the significant reduction of the difference between the analytical and the numerical calculations. Figure 2.3 shows another result of the projectile motion studies, now obtained with a step size of 0.1 s. The difference between the numerical and the analytical method is further reduced. The differences between the numerical method and the analytical method at the end of the simulation are shown in Table 2.1 for the step sizes used to generate Figs. 2.1, 2.2, and 2.3.

dt (s)	Difference (m)
10	4900
1	515
0.1	52

Table 2.1: Difference between the z coordinate numerical simulations at the end of the simulation and the analytical solution for different step sizes dt.

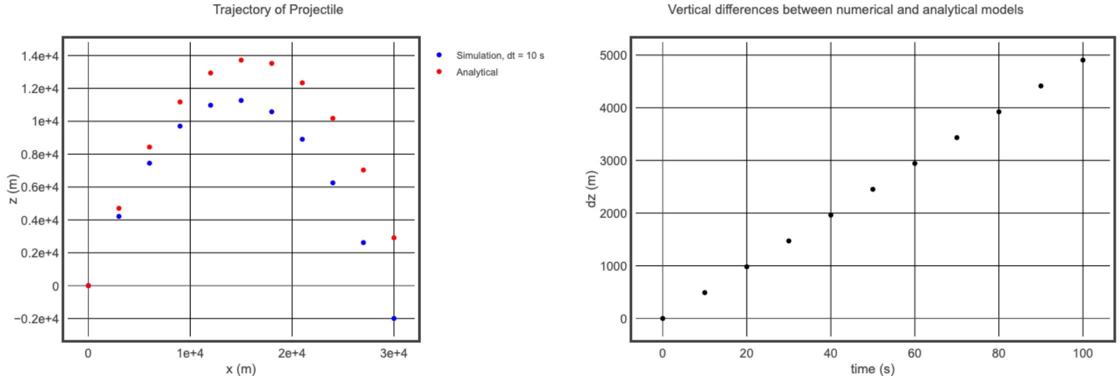


Figure 2.1: Results of numerical and analytical calculations of projectile motion with $dt = 10$ s. Left: vertical versus horizontal position of the projectile. Right: the difference in the vertical position of the projectile obtained from the analytical solution and the height of the projectile obtained with numerical simulations.

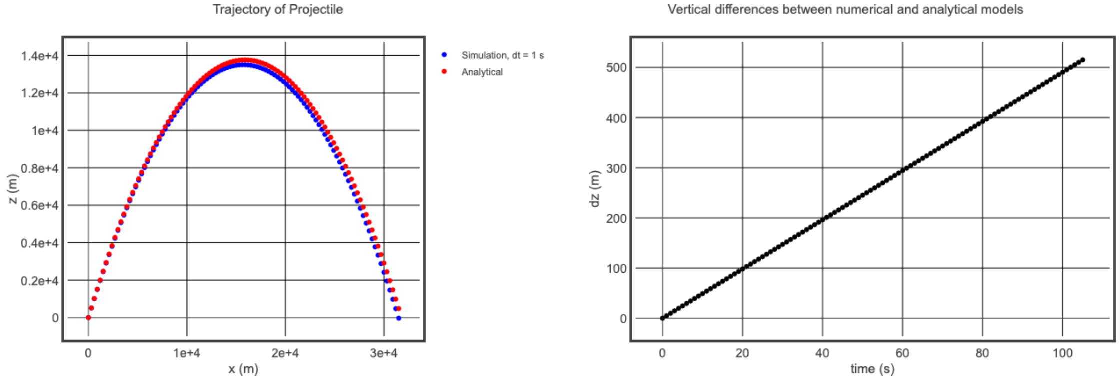


Figure 2.2: Results of numerical and analytical calculations of projectile motion with $dt = 1$ s. Left: vertical versus horizontal position of the projectile. Right: the difference in the vertical position of the projectile obtained from the analytical solution and the height of the projectile obtained with numerical simulations.

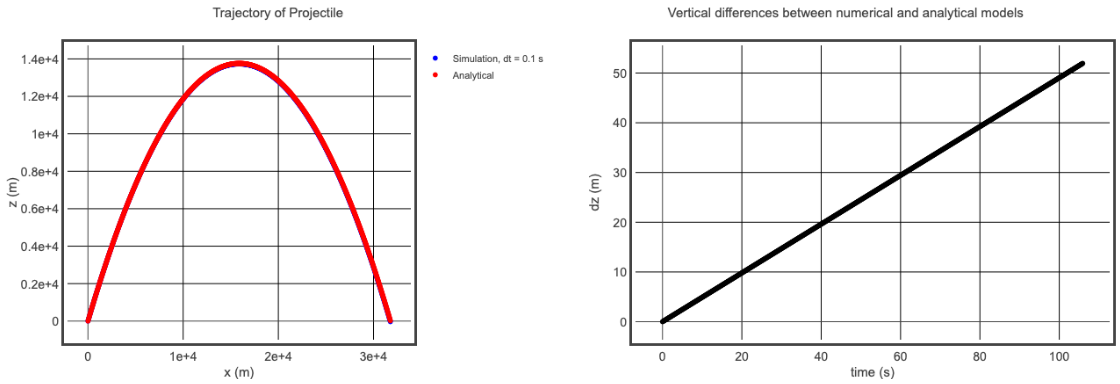


Figure 2.3: Results of numerical and analytical calculations of projectile motion with $dt = 0.1$ s. Left: vertical versus horizontal position of the projectile. Right: the difference in the vertical position of the projectile obtained from the analytical solution and the height of the projectile obtained with numerical simulations.

Now let us consider what happens when besides the gravitational force, there is a drag force acting on the particle. The drag force is usually proportional to the speed of the particle and directed in a direction opposite to the direction of motion. The net force on the particle is this equal to

$$F_x = -kmv_x \tag{2.9}$$

$$F_y = -kmv_y - mg \tag{2.10}$$

Since the force is known, we can determine the velocity of the particle:

$$v_x(t+dt) = v_x(t) + \frac{dt}{m}F_x = v_x(t) + \frac{dt}{m}(-kmv_x(t)) = (1-kdt)v_x(t) \tag{2.11}$$

$$v_y(t+dt) = v_y(t) + \frac{dt}{m}F_y = v_y(t) + \frac{dt}{m}\left(-kmv_y(t) - mg\right) = (1-kdt)v_y(t) - gdt \tag{2.12}$$

The position of the particle can be found once we have determined the velocity as function of time:

$$x(t+dt) = x(t) + v_x(t)dt \tag{2.13}$$

$$y(t+dt) = y(t) + v_y(t)dt \tag{2.14}$$

/noindent In principle, we can still solve this problem analytically (see Example 2.5 in the text book), but we will use the same numerical approach as we used in our study of projectile motion without drag to study the trajectory for different values of k. The results of a calculation for $k = 0.01$ and with time steps of 0.25 s is shown in Figure 4.

2.4 Conservation Laws

Several important conservation laws are a direct consequence of Newton's laws of motion. These conservation laws can significantly reduce the effort required to solve certain mechanics problems. In this Section we will briefly discuss the most important conservation laws that we will use in classical mechanics.

1. **The total linear momentum $\vec{p}$ of a particle is conserved when the total force on it is zero.**

   This law is a direct consequence of Newton's second law, which relates the change in the linear momentum of a particle to the force acting on it.

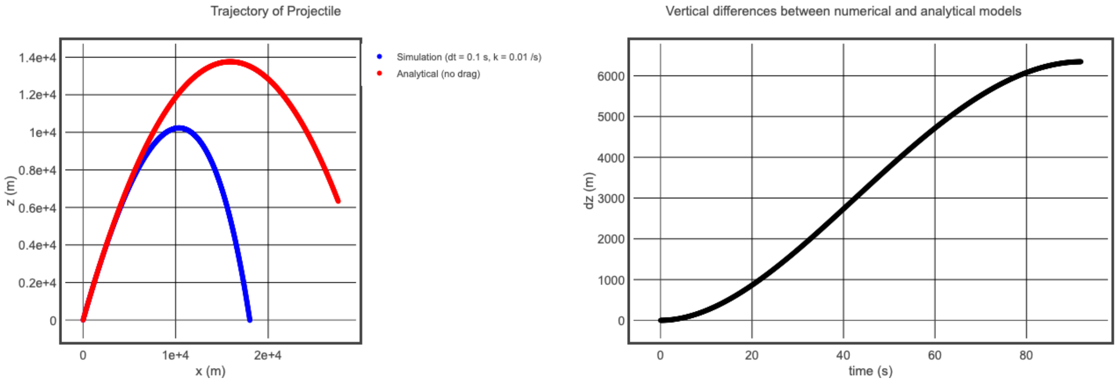


Figure 2.4: Results of numerical and analytical calculations of projectile motion with $dt = 0.1$ s and $k = 0.01$ s^{-1}. Left: vertical versus horizontal position of the projectile. Right: the difference in the vertical position of the projectile obtained from the analytical solution and the height of the projectile obtained with numerical simulations.

2. **The angular momentum $\vec{L}$ of a particle subject to no torque is conserved.**

   This law is a direct consequence of the definition of angular momentum and torque. In fact one can argue that torque was defined such that it is equal to the rate of change of the angular momentum ($d\vec{L}/dt$).

3. **The total energy E of a particle in a conservative force field is constant in time.**

   The total energy E is defined as the sum of the kinetic energy T and the potential energy U. The potential energy U is defined by the force field only to within a constant. It has no absolute meaning, and only differences in the potential energy are physically meaningful. The force field is conservative if the line integral of the force between two points is path independent. In this case, we can write the force as the gradient of a scalar function, and this scalar function is the potential energy U:

$$\vec{F} = -\vec{\nabla} U \tag{2.15}$$

   We can show that the rate of energy change dE/dt will be zero if the potential energy U does not depend explicitly on time ($\partial U/\partial t = 0$).

These three conservation laws are the most important conservation laws in classical mechanics and we will use them in many different applications.
An important application of one of our conservation laws is the prediction of motion based on a potential energy curve $U(x)$. Consider for example, the potential energy curve shown in Fig. 2.5. Since the total energy E is the sum of the potential energy U and the kinetic energy T, the total energy will always be larger or equal to the potential energy U. Looking at Fig. 2.5, we can immediately draw some important conclusions:

1. No particle can exist with a total energy E less than E_0.
2. A particle with energy E_1 can only be present between x_a and x_b.
3. A particle with energy E_4 can be present at any position.

Since the force is related to the derivative of the potential energy, the positions where the derivative is equal to 0 are the positions where the net force the on the particle is zero (these are the equilibrium positions).
Consider the potential energy in the vicinity of an equilibrium position, and assume we have chosen our coordinate system such that the equilibrium position corresponds to $x = 0$. We can expand the potential around the equilibrium point:

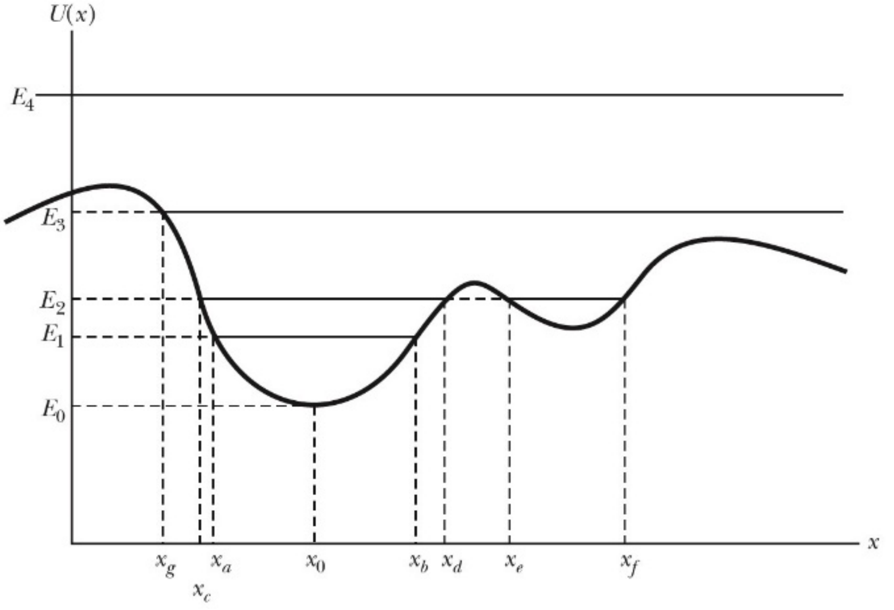


Figure 2.5: Potential energy $U(x)$ as function of position.

$$U(x) = U_0 + x\left(\frac{dU}{dx}\right)_0 + \frac{x^2}{2!}\left(\frac{d^2U}{dx^2}\right)_0 + \frac{x^3}{3!}\left(\frac{d^3U}{dx^3}\right)_0 + \ldots \tag{2.16}$$

Since at the equilibrium point, the slope of $U(x)$ is zero ($dU/dx = 0$) and since we can define the potential to be zero at this equilibrium point, we can rewrite the expansion of U as

$$U(x) = \frac{x^2}{2!}\left(\frac{d^2U}{dx^2}\right)_0 + \frac{x^3}{3!}\left(\frac{d^3U}{dx^3}\right)_0 + \ldots \tag{2.17}$$

For small displacements with respect to the equilibrium position, x is small. As a result, the first non-zero term in the expansion will dominate the expansion:

$$U(x) = \frac{x^2}{2!}\left(\frac{d^2U}{dx^2}\right)_0 \tag{2.18}$$

For the equilibrium to be **stable**, the potential energy on either side of the equilibrium point must be higher than the potential energy at the equilibrium point (which we defined to be 0). In order to achieve this we must require that

$$\left(\frac{d^2U}{dx^2}\right)_0 > 0 \tag{2.19}$$

If this condition is not satisfied, the equilibrium is an **unstable** equilibrium.

2.5 Limitations of Newtonian Mechanics

Newtonian mechanics can be used to describe many every-day macroscopic phenomena. When we study macroscopic phenomena, we can measure both the position and the linear momentum of objects of interest with great precision. However, when we start to study microscopic object we discover that we can no longer measure the position and the linear momentum with great precision. In fact, the actual measurement may influence the state of the system. In this regime, our capability of determining the position and the linear momentum of an object are limited by the Heisenberg uncertainty principle, which states that

$$\Delta x \Delta p \geq 10^{-34} \text{ Js} \tag{2.20}$$

If we measure the position with infinite precision, the uncertainty in the linear momentum approaches infinity. In this regime, Newtonian mechanics can no longer be used, and we need **quantum mechanics** to describe microscopic systems.

The limitations of Newtonian mechanics also appear when we study motion with velocities close to the speed of light. In this regime, we need the **theory of relativity**. One fundamental assumption in the theory of relativity is that the speed of light is constant, the same in each reference frame. This is clearly inconsistent with Newtonian mechanics, and the rules that govern transformations of position and velocity between coordinate systems.

Another limitation of Newtonian mechanics becomes obvious when we try to describe systems with large numbers of particles. Even if we know all of the details of the interaction between the particles, is becomes very difficult to predict the properties of the system by carrying out calculations involving the each individual interaction between all the particles. Such systems can be described by theory of **statistical mechanics**, which relates the properties of microscopic interactions to the average macroscopic properties of the system.

3 Oscillations

In this Chapter different types of oscillations will be discussed. A particle carrying out oscillatory motion, oscillates around a stable equilibrium position (note: if the equilibrium position was a position of unstable equilibrium, the particle would not return to its equilibrium position, and no oscillatory motion would result). We will not only focus on harmonic motion in one dimension but also consider motion in two- or three-dimensions. In addition, we will discuss damped motion and driven motion.

3.1 Simple-Harmonic Motion in One Dimension

Since we have the freedom to choose the origin of our coordinate system, we can choose it to coincide with the equilibrium position of our oscillator. The force that is responsible for the harmonic motion can be expanded around the equilibrium position:

$$F(x) = F(0) + x\left(\frac{dF}{dx}\right)_0 + \frac{x^2}{2!}\left(\frac{d^2F}{dx^2}\right)_0 + \frac{x^3}{3!}\left(\frac{d^3F}{dx^3}\right)_0 + ... \tag{3.1}$$

where F_0 is the force at $x = 0$. Since $x = 0$ is the equilibrium point, the force at this point must be 0. If the displacement x is small, we can ignore all terms involving x^2 and higher powers. The force can thus be approximated by

$$F(x) = x\left(\frac{dF}{dx}\right)_0 = -kx \tag{3.2}$$

Note: the constant k is positive since dF/dx must be negative if the equilibrium point is a point of stable equilibrium. A force that is proportional to $-kx$ is said to obey **Hooke's law**.
Since the force F is equal to ma, we can rewrite Eq.!3.2 as

$$\ddot{x} + \frac{k}{m}x = 0 \tag{3.3}$$

This equation is frequently rewritten as

$$\ddot{x} + {\omega_0}^2 x = 0 \tag{3.4}$$

where ω_0 is the angular frequency of the harmonic motion. The most general solution of this differential equation is $A\ \cos(\omega_0 t\ \text{-}\ \phi)$ or $A\ \sin(\omega_0 t\ \text{-}\ \phi)$. The amplitude and the phase angle must be determined to match the initial conditions (for example, the position and velocity at time $t = 0$ may have been specified). The total energy of the harmonic oscillator is constant, but the kinetic and potential energy components of the total energy are time dependent. The angular frequency of the harmonic motion is independent of the amplitude of the motion. Systems that have this property are called **isochronous systems**.
Note: keep in mind though that only for small oscillations the force might be proportional to the displacement. This means that only for small displacements, the motion will be simple harmonic; for larger displacements the angular frequency may become dependent on amplitude.

3.2 Simple-Harmonic Motion in Two Dimensions

When we discuss simple-harmonic motion in two dimensions we can always decompose the motion into two components, each directed along one of the two coordinate axes. We need to consider two special cases:

1. **The motion is a result of a single force, which obeys Hooke's law.** The force $\vec{F}$ can be written as:

   $$\vec{F} = -k\vec{r} \tag{3.5}$$

   Of course, one can argue that with the proper choice of coordinate system, this can be considered one-dimensional motion. However, the reason that the solution discussed in the previous section is not the general solution in the two-dimensional case is due to the fact that the initial conditions, such

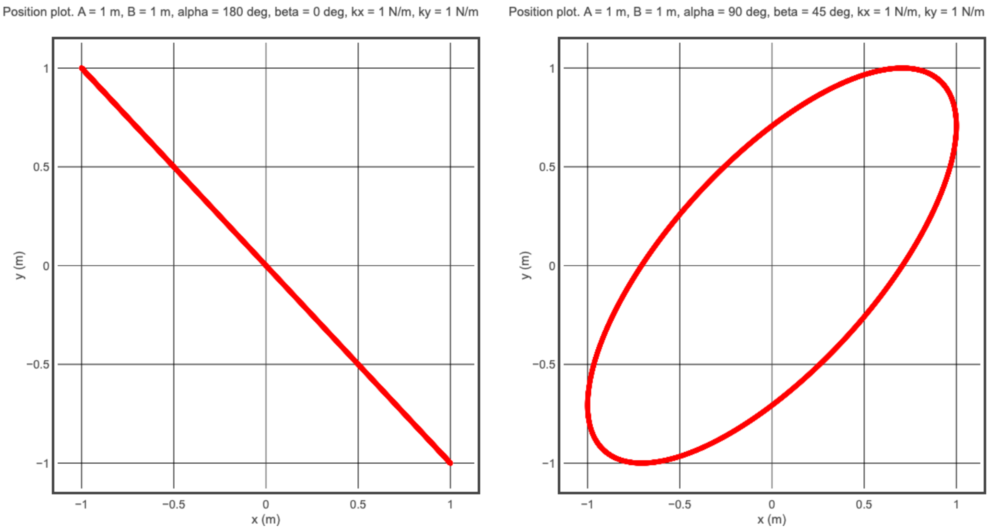


Figure 3.1: Two-dimensional motion of a simple-harmonic oscillator with $A = B = 1$ m, $\alpha = 180°$, $\beta = 0°$ (left) and $A = B = 1$ m, $\alpha = 90°$, $\beta = 45°$ (right). The angular frequencies of the motion in the x and y directions are taken to be the same.

as the velocity at time $t = 0$, do not have to be directed in the same direction as the position vector $\vec{r}$. If we decompose the position vector $\vec{r}$ into its components along the x and y axes, we get the following differential equations that must be satisfied by the oscillator:

$$\ddot{x} + {\omega_0}^2 x = 0 \tag{3.6}$$

$$\ddot{y} + {\omega_0}^2 y = 0 \tag{3.7}$$

The resulting motion will be the combined motion of two simple-harmonics oscillators, one along each axes, with the same angular frequency but different amplitudes and phase angles:

$$x(t) = A\cos(\omega_0 t - \alpha) \tag{3.8}$$

$$y(t) = B\cos(\omega_0 t - \beta) \tag{3.9}$$

Although the equations of motion shown in Eqs.3.8 and 3.9 are simple, the easiest way to examine this type of motion is by actually looking at plots of for example, the trajectory carried out by the particle. You can use Phy235-SimpleHarmonicMotion in the public glowscript folder to study the motion of this two-dimensional oscillator. Examples of trajectories for $A = B = 1$ m, $\alpha = 180°$, $\beta = 0°$ and $A = B = 1$ m, $\alpha = 90°$, $\beta = 45°$ are shown in Fig. 3.1.

Since the angular frequency of the motion in the x direction is the same as the angular frequency of the motion in the y direction, the trajectory of the particle will be a closed trajectory (that is, after a given period T, the particle will return to the same position and have the same velocity and acceleration; the motion is truly periodic).

2. **The motion is a result of a several forces, which each obey Hooke's law.**

   Now consider what happens if we have several forces acting on the particle. Since each of these forces may have a different force constant, the angular frequency of the motion along the x axis may be different from the angular frequency of the motion along the y axis:

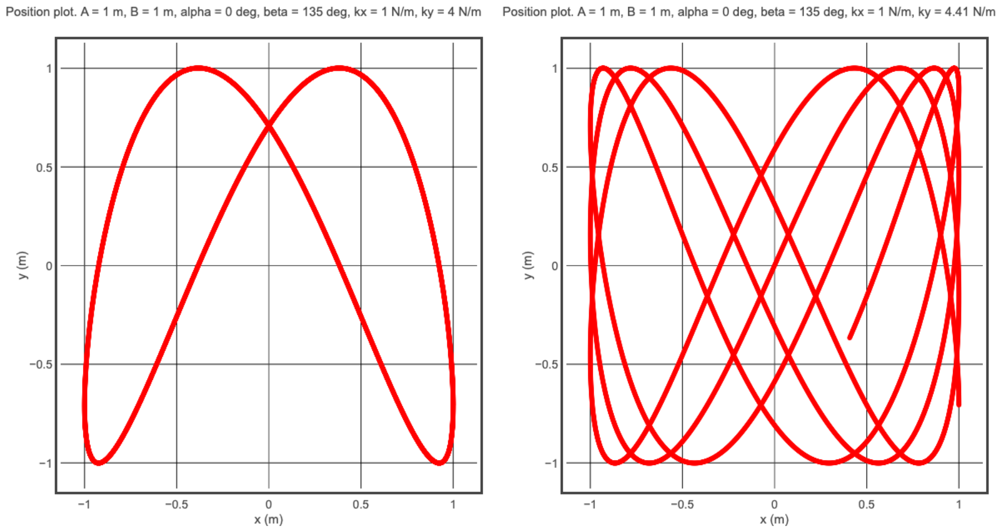


Figure 3.2: Lissajous curve for $A = B = 1$ m, $\alpha = 0°$, $\beta = 135°$, $\omega_A = 1$ rad/s, $\omega_B = 2$ rad/s (left) and $A = B = 1$ m, $\alpha = 0°$, $\beta = 135°$, $\omega_A = 1$ rad/s, $\omega_B = 2.1$ rad/s (right).

$$\ddot{x} + {\omega_A}^2 x = 0 \tag{3.10}$$

$$\ddot{y} + {\omega_B}^2 y = 0 \tag{3.11}$$

The corresponding x and y positions as function of time are

$$x(t) = A\cos(\omega_A t - \alpha) \tag{3.12}$$

$$y(t) = B\sin(\omega_B t - \beta) \tag{3.13}$$

The biggest difference between the case of multiple forces and the case of a single force is that in the former case there is no guarantee that the trajectory is closed (except when ω_A/ω_B is a rational fraction), while in the latter case every trajectory is closed, independent of the initial conditions.

The trajectory described by the particle is called a Lissajous curve. Examples of such curves for two slightly different sets of parameters are shown in Fig.3.2 for $A = B = 1$ m, $\alpha = 0°$, $\beta = 135°$, $\omega_A = 1$ rad/s, $\omega_B = 2$ rad/s (left) and $A = B = 1$ m, $\alpha = 0°$, $\beta = 135°$, $\omega_A = 1$ rad/s, $\omega_B = 2.1$ rad/s (right).

3.3 Phase Diagrams

Although the trajectory of the particle in real space is one way to visualize the information of the motion of the oscillators, in general it does not provide information about important parameters such as the total energy of the system. More detailed information is provided by phase diagrams; they show simultaneous information about the position and the velocity of the particle (which is the information that is required to uniquely specify the motion of the simple-harmonic oscillator). For example, a phase diagram for a one-dimensional oscillator is a two-dimensional figure showing velocity versus position. Figure 3.3 shows a phase diagram for a one-dimensional simple-harmonic oscillator; it shows three phase paths, corresponding to three different total energies. A few important observations about phase diagrams can be made:

1. **Two phase paths can not cross.** If they would cross at a particular point, then the total energy at that point would not be defined (it would have two possible values).

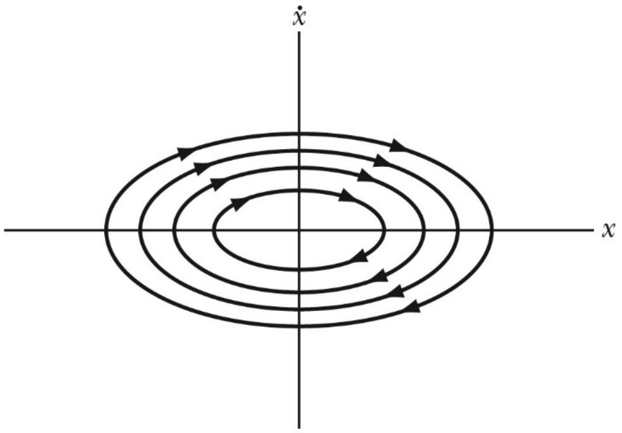


Figure 3.3: Phase diagram for a one-dimensional simple-harmonic oscillator

2. **The phase paths will be executed in a clock-wise direction.** For example, in the upper right corner of the phase diagram, the velocity is positive. This implies that x must be increasing. The x coordinate will continue to increase until the velocity becomes equal to zero. When the velocity becomes negative, the x coordinate will start to decrease.

3.4 Solving Second-Order Differential Equations

The second-order differential equations that we have discussed in the previous sections are simple equations that can be solved analytically with little effort. Once we start including damping and/or driving forces, the equations become more complicated, and we need to discuss in more detail how we can solve these equations.

Second-order differential equations have the following form:

$$\ddot{y} + a\dot{y} + by = f(x) \tag{3.14}$$

where a and b are constants. When $f(x)) = 0$, the equation is called a **homogeneous equation**; otherwise it is called an **inhomogeneous equation**. Any solution of this equation can be rewritten as a linear superposition of any two linearly independent solutions of this equation. Any solution will have two parameters that need to be adjusted to match the initial conditions.

We start by first looking at the homogeneous equation. Consider the following solution of this equation:

$$y(x) = e^{rx} \tag{3.15}$$

If we substitute this solution into the homogeneous equation (Eq. 3.14) we get

$$r^2 + ar + b = 0 \tag{3.16}$$

In general, there are two possible values of r:

$$r_\pm = -\frac{a}{2} \pm \frac{1}{2}\sqrt{a^2 - 4b} \tag{3.17}$$

If $a \neq b$ then we can write the general solution to the homogeneous differential equation as

$$y(x) = c_1 e^{\left(-\frac{a}{2} + \frac{1}{2}\sqrt{a^2-4b}\right)x} + c_2 e^{\left(-\frac{a}{2} - \frac{1}{2}\sqrt{a^2-4b}\right)x} \tag{3.18}$$

This is the most general solution to the homogeneous differential equation since its two terms are linearly independent.

If $a = b$ then there is only one solution for r and there must be another solution of the differential equation. Consider the function xe^{rx}, where $r = -a/2$. This is a solution of the homogeneous differential equation, which we can verify by substituting it into the equation:

$$\frac{d^2}{dx^2}(xe^{rx}) + a\frac{d}{dx}(xe^{rx}) + b(xe^{rx}) = (2r+a)e^{rx} + \left(r^2 + ar + b\right)xe^{rx} = 0 \tag{3.19}$$

Since the functions xe^{rx} and e^{rx} are linearly independent, the most general solution can be written as

$$y(x) = c_1 e^{\left(-\frac{a}{2}\right)x} + c_2 x e^{\left(-\frac{a}{2}\right)x} \tag{3.20}$$

If $a < 4b$, the solutions for r are complex numbers:

$$r_\pm = \alpha \pm i\beta \tag{3.21}$$

and the general solution can be rewritten as

$$y(x) = e^{\alpha x}(c_1 \cos\beta x + c_2 \sin\beta x) \tag{3.22}$$

Now we continue and consider the inhomogeneous equation

$$\ddot{y} + a\dot{y} + by = f(x) \tag{3.23}$$

Consider that u is the general solution of the corresponding homogenous solution (u is called the **complementary function)** and v is a solution of the inhomogeneous equation (v is called the a **particular solution)**. This requires that

$$\ddot{u} + a\dot{u} + bu = 0 \tag{3.24}$$

and

$$\ddot{v} + a\dot{v} + bv = f(x) \tag{3.25}$$

The function $u + v$ is a solution of the inhomogeneous solution, and since it contains two constants (since u is the general solution of the homogeneous solution), it must be the general solution of the inhomogeneous equation.

3.5 Damped Oscillations

A damped oscillator has a restoring force that satisfies Hooke's law and a damping force that may be a function of velocity. The differential equation that describes the damped motion is

$$m\ddot{x} + b\dot{x} + kx = 0 \tag{3.26}$$

or

$$\ddot{x} + \frac{b}{m}\dot{x} + \frac{k}{m}x = \ddot{x} + 2\beta\dot{x} + {\omega_0}^2 x = 0 \tag{3.27}$$

where β is the damping parameter. One solution of this equation is $e^{\gamma t}$ where

$$\gamma = \frac{-2\beta \pm \sqrt{4\beta^2 - 4\omega_0^2}}{2} = -\beta \pm \sqrt{\beta^2 - \omega_0^2} \tag{3.28}$$

Since there are two possible values of γ, the general solution of the differential equation can be written as

$$x(t) = e^{-\beta t}\left\{A_1 e^{\sqrt{\beta^2 - {\omega_0}^2}t} + A_2 e^{-\sqrt{\beta^2 - {\omega_0}^2}t}\right\} \tag{3.29}$$

The corresponding velocity is equal to

$$v(t) = \left\{ \sqrt{\beta^2 - {\omega_0}^2} - \beta \right\} A_1 e^{-\beta t} e^{\sqrt{\beta^2 - {\omega_0}^2} t} - \left\{ \sqrt{\beta^2 - {\omega_0}^2} + \beta \right\} A_2 e^{-\beta t} e^{-\sqrt{\beta^2 - {\omega_0}^2} t} \tag{3.30}$$

There are three distinct modes of oscillations described by this general solution. To describe each mode, we assume that at time $t = 0$, the displacement is x_0 and the velocity is v_0. The phase diagrams for these examples can be explored using the glowscript program Phy235-3D-DampedHarmonicMotion in https://www.glowscript.org/#/user/wolfs/folder/Public/.

1. **Critical Damping:** ${\omega_0}^2 = \beta^2$.

   In this case, Eq. 3.28 has only one value and we thus need a second solution to Eq. 3.26. This second solution is $te^{-\beta t}$. The most general solution Eq. 3.26 is thus equal to

   $$x(t) = A_1 e^{-\beta t} + A_2 t e^{-\beta t} \tag{3.31}$$

   The velocity of mass m is given by

   $$v(t) = -\beta A_1 e^{-\beta t} + A_2 \left(e^{-\beta t} - \beta t e^{-\beta t} \right) \tag{3.32}$$

   In order to satisfy the boundary conditions, we must require that

   $$x(0) = A_1 = x_0 \tag{3.33}$$

   $$v(0) = -\beta A_1 + A_2 = -\beta x_0 + A_2 = v_0 \tag{3.34}$$

   or

   $$A_1 = x_0 \tag{3.35}$$

   $$A_2 = v_0 + \beta x_0 \tag{3.36}$$

   An example of a phase diagram for critical damping is shown in Fig. 3.4(left). The total energy as function of time for critical damping is shown in Fig. 3.5(top).

2. **Over damping:** ${\omega_0}^2 < \beta^{\,2}$.

   In this case, the position and velocity are given by

   $$x(t) = e^{-\beta t} \left\{ A_1 e^{\sqrt{\beta^2 - {\omega_0}^2} t} + A_2 e^{-\sqrt{\beta^2 - {\omega_0}^2} t} \right\} \tag{3.37}$$

   and

   $$v(t) = e^{-\beta t} \left\{ A_1 \left\{ \sqrt{\beta^2 - {\omega_0}^2} - \beta \right\} e^{\sqrt{\beta^2 - {\omega_0}^2} t} - A_2 \left\{ \sqrt{\beta^2 - {\omega_0}^2} + \beta \right\} e^{-\sqrt{\beta^2 - {\omega_0}^2} t} \right\} \tag{3.38}$$

   To match the boundary conditions, we must require that

   $$x(0) = A_1 + A_2 = x_0 \tag{3.39}$$

   $$v(0) = A_1 \left\{ \sqrt{\beta^2 - {\omega_0}^2} - \beta \right\} - A_2 \left\{ \sqrt{\beta^2 - {\omega_0}^2} + \beta \right\} = v_0 \tag{3.40}$$

   or

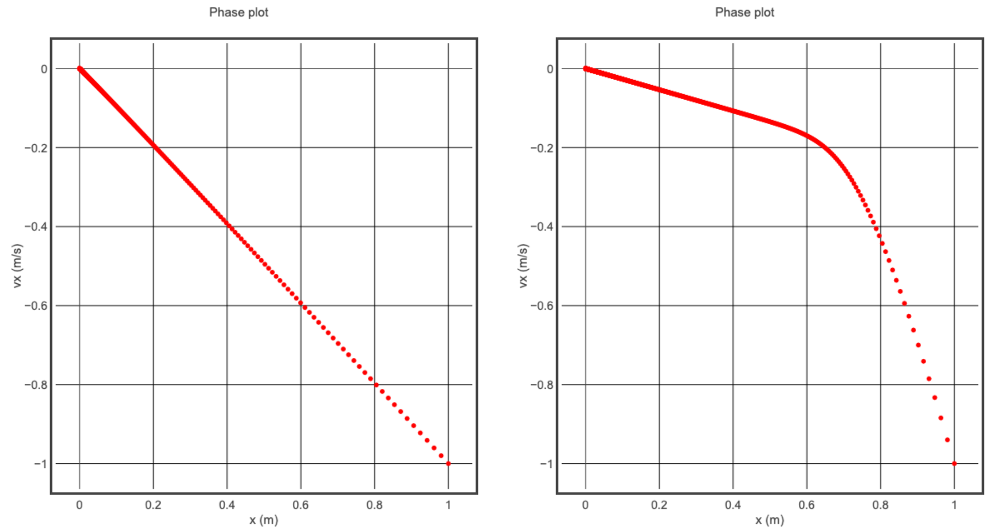


Figure 3.4: Phase diagram of critically damped harmonic motion with ω_0 = 1 rad/s and β = 1 Ns/m (left) and over damped harmonic motion with ω_0 = 1 rad/s, and β = 2 Ns/m (right). In these simulations, the initial position of the ball is (1 m, 0 m, 0 m) and the initial velocity is (-1 m/s, 0 m/s, 0 m/s)

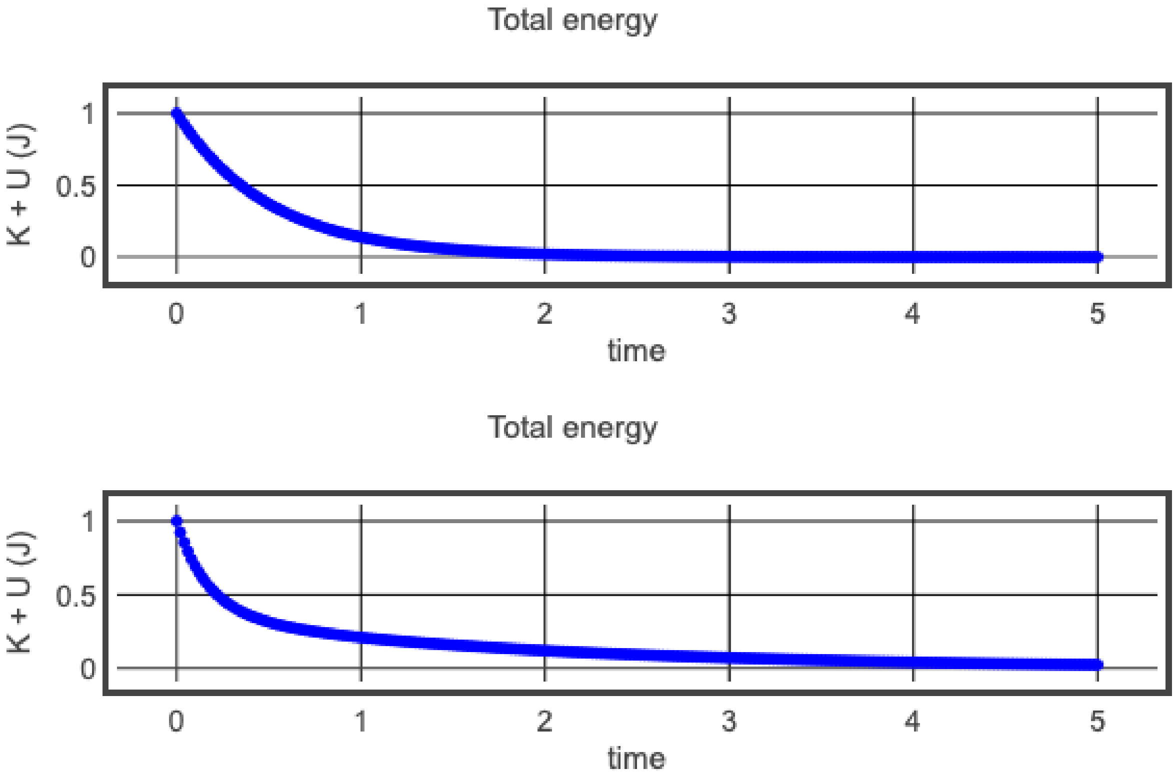


Figure 3.5: Total energy as function of time of critically damped harmonic motion with ω_0 = 1 rad/s and β = 1 Ns/m (top) and over damped harmonic motion with ω_0 = 1 rad/s, and β = 2 Ns/m (bottom). In these simulations, the initial position of the ball is (1 m,0 m,0 m) and the initial velocity is (-1 m/s, 0 m/s, 0 m/s)

$$A_1 = \frac{x_0 \left\{ \sqrt{\beta^2 - {\omega_0}^2} + \beta \right\} + v_0}{2\sqrt{\beta^2 - {\omega_0}^2}} \tag{3.41}$$

$$A_2 = \frac{x_0 \left\{ \sqrt{\beta^2 - {\omega_0}^2} - \beta \right\} - v_0}{2\sqrt{\beta^2 - {\omega_0}^2}} \tag{3.42}$$

An example of a phase diagram for over damping is shown in Fig. 3.4(right). The total energy as function of time for over damping is shown in Fig. 3.5(bottom).

3. **Under damping:** ${\omega_0}^2 > \beta^{\,2}$.

   In this case, the values of γ obtained from Eq. 3.28 are complex numbers:

$$\gamma = -\beta \pm i\sqrt{\omega_0^2 - \beta^2} \tag{3.43}$$

   The position and velocity of mass m can be written as

$$x(t) = e^{-\beta t}\left\{ A_1 \cos\left(\sqrt{{\omega_0}^2 - \beta^2}\, t \right) + A_2 \sin\left(\sqrt{{\omega_0}^2 - \beta^2}\, t \right) \right\} \tag{3.44}$$

   and

$$v(t) = e^{-\beta t}\left\{ \left(-\beta A_1 + \sqrt{{\omega_0}^2 - \beta^2} A_2 \right) \cos\left(\sqrt{{\omega_0}^2 - \beta^2}\, t \right) - \left(\beta A_2 + \sqrt{{\omega_0}^2 - \beta^2} \right) \sin\left(\sqrt{{\omega_0}^2 - \beta^2}\, t \right) \right\} \tag{3.45}$$

   To satisfy the boundary conditions, we must require that

$$x(0) = A_1 = x_0 \tag{3.46}$$

$$v(0) = -\beta A_1 + \sqrt{{\omega_0}^2 - \beta^2} A_2 = v_0 \tag{3.47}$$

   or

$$A_1 = x_0 \tag{3.48}$$

$$A_2 = \frac{\beta A_1 + v_0}{\sqrt{{\omega_0}^2 - \beta^2}} \tag{3.49}$$

   Examples of phase diagrams for under damping are shown in Fig. 3.6. The total energy as function of time for critical damping is shown in Fig. 3.7.

A comparison of the displacement as function of time for the three types of damping is shown in Fig. 3.8. For a given set of initial conditions (same displacement and velocity) we see that the critically damped oscillator approaches the equilibrium position at a rate that is higher than either the under damped or the over damped oscillator.

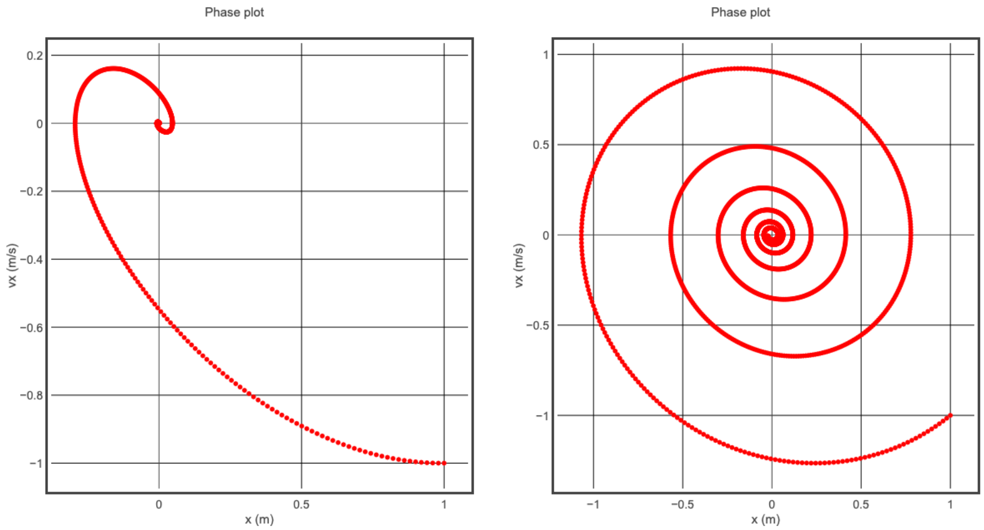


Figure 3.6: Phase diagram of under damped harmonic motion with $\omega_0 = 1$ rad/s and $\beta = 0.5$ Ns/m (left) and $\omega_0 = 1$ rad/s and $\beta = 0.1$ Ns/m (right). In these simulations, the initial position of the ball is (1 m, 0 m, 0 m) and the initial velocity is (-1 m/s, 0 m/s, 0 m/s)

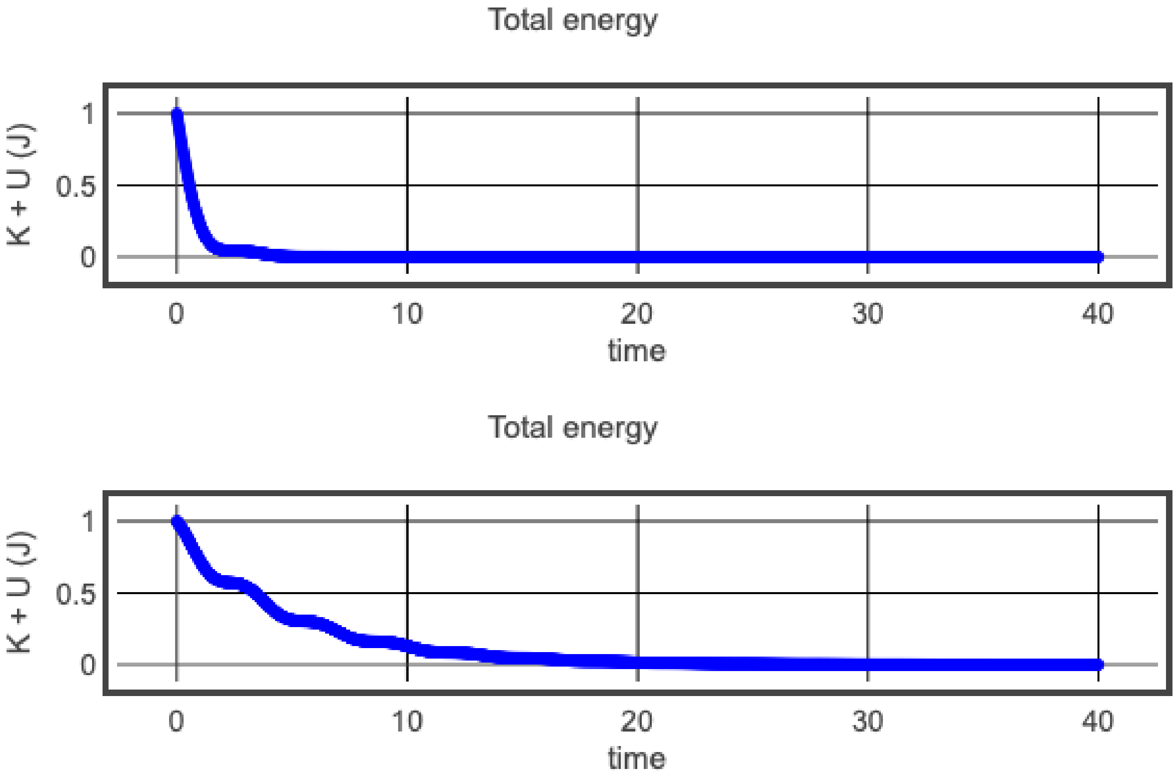


Figure 3.7: Total energy as function of time of under damped harmonic motion with $\omega_0 = 1$ rad/s, and $\beta = 0.5$ Ns/m (top) and $\omega_0 = 1$ rad/s, and $\beta = 0.1$ Ns/m (bottom). In these simulations, the initial position of the ball is (1 m, 0 m, 0 m) and the initial velocity is (-1 m/s, 0 m/s, 0 m/s)

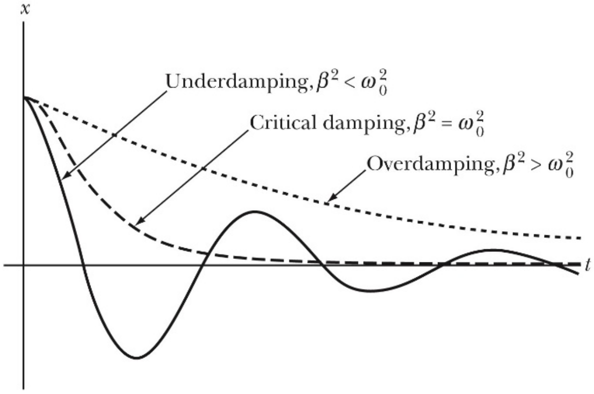


Figure 3.8: Displacement as function of time for a one-dimensional damped oscillator.

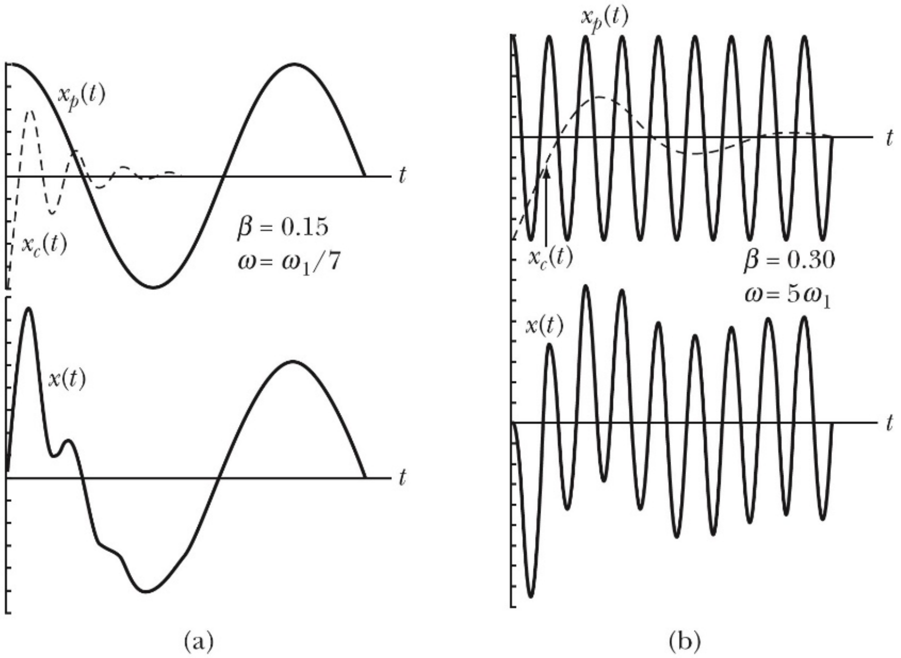


Figure 3.9: Examples of sinusoidal driven oscillatory motion with damping.

3.6 Driven Motion

We will start our discussion of driven motion by considering an external driving force that varies in a harmonic fashion as function of time. The equation of motion for an oscillator exposed to this type of driving force is given by

$$\ddot{x} + 2\beta\dot{x} + {\omega_0}^2 x = A\cos(\omega t) \tag{3.50}$$

The complementary solution has already been discussed previously, see Eq. 3.29, and is given by

$$x_c(t) = e^{-\beta t}\left\{A_1 e^{\sqrt{\beta^2-{\omega_0}^2}t} + A_2 e^{-\sqrt{\beta^2-{\omega_0}^2}t}\right\} \tag{3.51}$$

When $t \to \infty$, $x_c \to 0$. A particular solution is

$$x_p(t) = \frac{A}{\sqrt{({\omega_0}^2-\omega^2)^2 + 4\omega^2\beta^2}}\cos(\omega t - \delta) \tag{3.52}$$

where

$$\delta = \tan^{-1}\left(\frac{2\omega\beta}{{\omega_0}^2-\omega^2}\right) \tag{3.53}$$

This can be verified by substituting this solution in the original differential equation (Eq. 3.50).

The complementary solution has an amplitude that decreases exponentially as a function of time, while the amplitude of the particular solution is independent of time. As a result, for long times, the motion will be dominated by the particular solution, and this solution represents the steady-state solution. The transient effects may be dominated by the complementary solution. Figure 3.9 shows examples of complementary functions and particular solutions for a driven oscillator with damping.

The amplitude of the steady-state solution of the driven harmonic oscillator is a function of the angular frequency ω of the driving force and the "natural" angular frequency of the oscillator. There are three different resonance frequencies for the system:

1. **The amplitude resonance frequency.**

   This is the frequency for which the amplitude of the oscillator in its steady-state condition has a maximum. This requires that

   $$\frac{d}{d\omega}\left(\frac{A}{\sqrt{({\omega_0}^2-\omega^2)^2+4\omega^2\beta^2}}\right) = -\frac{-2\omega\left({\omega_0}^2-\omega^2\right)+4\omega\beta^2}{\left(({\omega_0}^2-\omega^2)^2+4\omega^2\beta^2\right)^{3/2}}A = 0 \tag{3.54}$$

   or

   $$\omega\left(-{\omega_0}^2+\omega^2+2\beta^2\right) = 0 \tag{3.55}$$

   The amplitude has a maximum when the driving frequency is equal to

   $$\omega_R = \sqrt{{\omega_0}^2 - 2\beta^2} \tag{3.56}$$

   When $\beta \neq 0$, the amplitude of the particular solution, shown in Eq. 3.52, will not go to inf when $\omega = \omega_R$.

2. **The kinetic energy resonance frequency.**

   The kinetic energy of the oscillator in its steady-state condition is equal to

$$T = \frac{1}{2}m\dot{x}^2 = \frac{1}{2}m\left(\frac{-A\omega}{\sqrt{\left({\omega_0}^2-\omega^2\right)^2+4\omega^2\beta^2}}\sin\left(\omega t-\delta\right)\right)^2 = \frac{1}{2}m\frac{A^2\omega^2}{\left({\omega_0}^2-\omega^2\right)^2+4\omega^2\beta^2}\sin^2\left(\omega t-\delta\right) \quad (3.57)$$

The average kinetic energy, averaged over one period, is equal to

$$\langle T\rangle = \frac{1}{2}m\frac{A^2\omega^2}{\left({\omega_0}^2-\omega^2\right)^2+4\omega^2\beta^2}\left\langle\sin^2\left(\omega t-\delta\right)\right\rangle = \frac{1}{4}m\frac{A^2\omega^2}{\left({\omega_0}^2-\omega^2\right)^2+4\omega^2\beta^2} \quad (3.58)$$

The average kinetic energy has a maximum when

$$\begin{aligned}\frac{d\langle T\rangle}{d\omega} &= \frac{1}{4}m\left\{\frac{2A^2\omega}{\left({\omega_0}^2-\omega^2\right)^2+4\omega^2\beta^2} - \frac{A^2\omega^2\left(8\omega\beta^2-4\omega\left({\omega_0}^2-\omega^2\right)\right)}{\left(\left({\omega_0}^2-\omega^2\right)^2+4\omega^2\beta^2\right)^2}\right\} = \\ &= \frac{1}{4}m\left\{\frac{2A^2\omega^2\left({\omega_0}^2-\omega^2\right)\left({\omega_0}^2+\omega^2\right)}{\left(\left({\omega_0}^2-\omega^2\right)^2+4\omega^2\beta^2\right)^2}\right\} = 0\end{aligned} \quad (3.59)$$

This happens when the driving frequency ω equal to natural frequency ω_0.

3. **The potential energy resonance frequency.**

   Since the potential energy is proportional the displacement squared, the potential energy will have a maximum when the displacement has a maximum. This resonance frequency of the potential energy will thus be the same as the amplitude resonance frequency.

Although in general the driving force will not be a simple harmonic function, the discussion in this Section is useful since any arbitrary function can be expanded in a series of harmonic functions. Consider a driving force $F(t)$ and assume the driving force has a period t. Consider that the force F can be expanded in the following way in terms of simple harmonic functions:

$$F(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}\left(a_n\cos(n\omega t) + b_n\sin(n\omega t)\right) \quad (3.60)$$

where

$$\omega = \frac{2\pi}{\tau} \quad (3.61)$$

Consider the following integrals:

$$\int_0^{\tau} F(t')\cos(n\omega t')dt' = \int_0^{\tau} a_n\cos^2(n\omega t')dt' = \frac{a_n}{n\varpi}\int_0^{n\omega\tau}\cos^2(x)dx = a_n\frac{\tau}{2} \quad (3.62)$$

$$\int_0^{\tau} F(t')\sin(n\omega t')dt' = \int_0^{\tau} b_n\sin^2(n\omega t')dt' = \frac{b_n}{n\varpi}\int_0^{n\omega\tau}\sin^2(x)dx = b_n\frac{\tau}{2} \quad (3.63)$$

They can be used to determine the coefficients a_n and b_n:

$$a_n = \frac{2}{\tau}\int_0^{\tau} F(t')\cos(n\omega t')dt' \quad (3.64)$$

$$b_n = \frac{2}{\tau}\int_0^{\tau} F(t')\sin(n\omega t')dt' \quad (3.65)$$

When we apply the equation of motion with force $F(t)$ to the oscillator, we need to solve the following differential equation to determine the motion of the oscillator:

$$\ddot{x} + 2\beta\dot{x} + {\omega_0}^2 x = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos n\omega t + b_n \sin n\omega t) \tag{3.66}$$

In order to solve this equation, we rely on the principle of superposition. Consider the function x_1 and x_2 which satisfy the following differential equations:

$$\ddot{x}_1 + 2\beta\dot{x}_1 + {\omega_0}^2 x_1 = F_1 \tag{3.67}$$

$$\ddot{x}_2 + 2\beta\dot{x}_2 + {\omega_0}^2 x_2 = F_2 \tag{3.68}$$

Adding these two equations, we obtain the following relation:

$$(\ddot{x}_1 + \ddot{x}_2) + 2\beta(\dot{x}_1 + \dot{x}_2) + {\omega_0}^2(x_1 + x_2) = F_1 + F_2 \tag{3.69}$$

We thus see that $(x_1 + x_2)$ is a solution of the differential equation with a driving force of $(F_1 + F_2)$. We can thus solve the differential equation for the driving force $F(t)$ if we solve each of the following differential equations:

$$\ddot{x} + 2\beta\dot{x} + {\omega_0}^2 x = \frac{1}{2} \tag{3.70}$$

$$\ddot{x} + 2\beta\dot{x} + {\omega_0}^2 x = \cos n\omega t = \cos\omega_n t \tag{3.71}$$

$$\ddot{x} + 2\beta\dot{x} + {\omega_0}^2 x = \sin n\omega t = \sin\omega_n t \tag{3.72}$$

The solution to Eq. 3.70 is

$$x = \frac{1}{2{\omega_0}^2} \tag{3.73}$$

The solution to Eq. 3.71 is

$$x(t) = \frac{A}{\sqrt{({\omega_0}^2 - {\omega_n}^2)^2 + 4{\omega_n}^2\beta^2}}\cos(\omega_n t - \delta_n) \tag{3.74}$$

where

$$\delta_n = \tan^{-1}\left(\frac{2\omega_n\beta}{{\omega_0}^2 - {\omega_n}^2}\right) \tag{3.75}$$

The solution of Eq. 3.72 is

$$x(t) = \frac{B}{\sqrt{({\omega_0}^2 - {\omega_n}^2)^2 + 4{\omega_n}^2\beta^2}}\sin(\omega_n t - \delta_n) \tag{3.76}$$

where

$$\delta_n = \tan^{-1}\left(\frac{2\omega_n\beta}{{\omega_0}^2 - {\omega_n}^2}\right) \tag{3.77}$$

The solution of the following differential equation

$$\ddot{x} + 2\beta\dot{x} + {\omega_0}^2 x = \frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos n\omega t + b_n \sin n\omega t) \tag{3.78}$$

is thus equal to

$$x(t) = \frac{a_0}{2{\omega_0}^2} + \sum_{n=1}^{\infty} \frac{1}{\sqrt{\left({\omega_0}^2 - {\omega_n}^2\right)^2 + 4{\omega_n}^2\beta^2}} \left(a_n \cos\left(\omega_n t - \delta_n\right) + b_n \sin\left(\omega_n t - \delta_n\right)\right) \tag{3.79}$$

where

$$\delta_n = \tan^{-1}\left(\frac{2\omega_n\beta}{{\omega_0}^2 - {\omega_n}^2}\right) \tag{3.80}$$

and

$$\omega_n = n\omega \tag{3.81}$$

4 Non-Linear Oscillations and Chaos

4.1 Non-Linear Differential Equations

Up to now, we have considered differential equations with terms that are proportional to the acceleration, the velocity, and the position:

$$m\ddot{x} + f(\dot{x}) + g(x) = h(t) \tag{4.1}$$

where $f(\dot{x})$ is proportional to $\dot{x}$ and $g(x)$ is proportional to x. Such an equation is called a **linear differential equation**. Although many important applications can be described in terms of a linear differential equation, many other applications require the inclusion of terms that are non linear, e.g. $\dot{x}^2$ and x^3. In those cases, the system is called **non-linear**. Differential equations can become non-linear if for example:

1. The restoring force is not proportional to the displacement x.
2. The damping force is not proportional to the velocity v.

An example of the first case is the pendulum. The pendulum will exhibit simple-harmonic motion if the angular displacement is small. In that case, the restoring force, which is proportional to $\sin\theta$, is approximately proportional to θ. However, if the angular displacement is large, $\sin\theta$ is no longer proportional to θ, and the system will no longer exhibit simple harmonic motion. Systems in which the restoring force is non-linear can be divided in to two groups:

1. **Systems for which the force is symmetric around the equilibrium position.**

   Examples of such forces are shown in Fig. 4.1. The non-linear behavior, which is added to the linear behavior, must be proportional to e.g. x^3 or x^5 since terms proportional to x^2 would reduce the magnitude of the restoring force on one side, while increasing its magnitude on the opposite side:

   $$F(x) = -kx + \varepsilon x^3 \tag{4.2}$$

   If $\varepsilon > 0$, the magnitude of the force is decreased compared to the linear situation. Such a system is called a **soft system**. If $\varepsilon < 0$, the magnitude of the force is increased compared to the linear situation. Such a system is called a **hard system**.

2. **Systems for which the force is asymmetric around the equilibrium position.**

   Examples of such forces are shown in Fig. 4.2. The non-linear behavior, which is added to the linear behavior, must be proportional to e.g. x^2 or x^4 since terms proportional to x^3 would reduce the magnitude of the restoring force on both sides:

   $$F(x) = -kx + \lambda x^2 \tag{4.3}$$

   If $\lambda > 0$, the force is increased compared to the linear situation. If $\lambda < 0$, the force is decreased compared to the linear situation. In both cases, the system is hard on one side of the equilibrium position, and soft on the opposing side.

4.2 Phase Diagrams for Non-Linear Systems

If the potential energy of the linear system is known, it is easy to construct the phase diagram. The procedure used to construct the phase diagram relies on conservation of energy:

$$E = U(x) + \frac{1}{2}m\dot{x}^2 \tag{4.4}$$

or

$$\dot{x} = \frac{2}{m}\sqrt{E - U(x)} \tag{4.5}$$

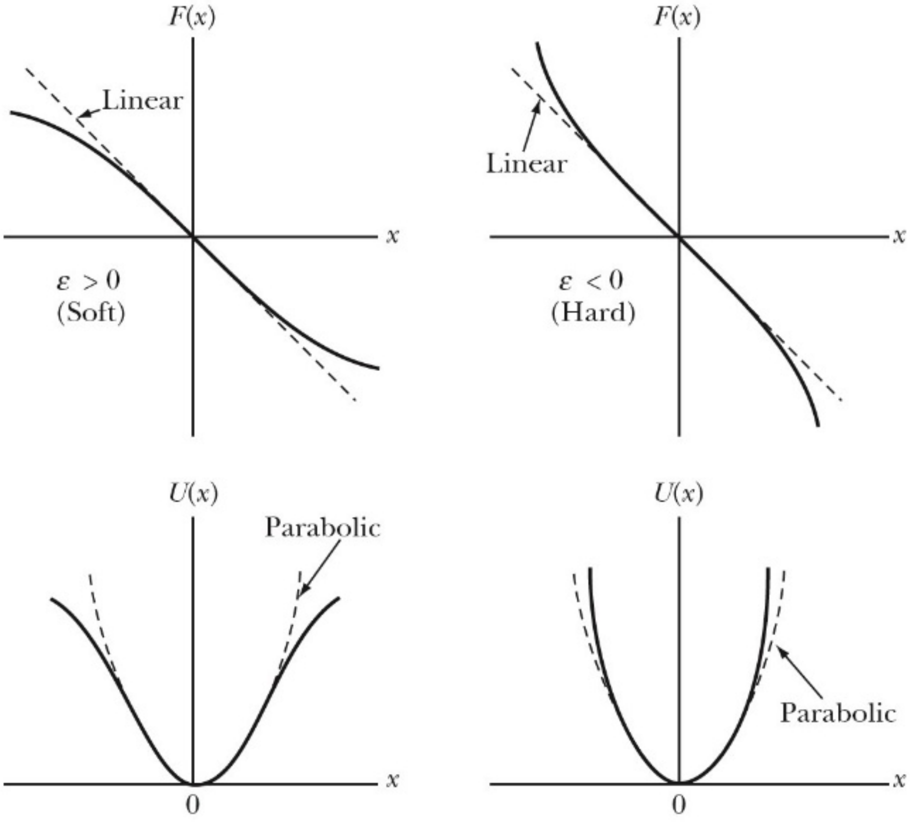


Figure 4.1: Examples of forces that are symmetric around the equilibrium position and their corresponding potential energy distributions.

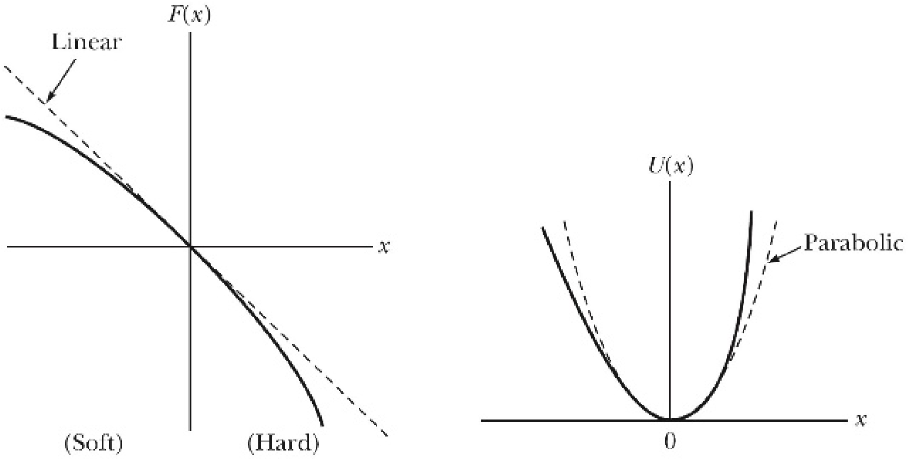


Figure 4.2: An example of a force that is asymmetric around the equilibrium position and its corresponding potential energy distribution.

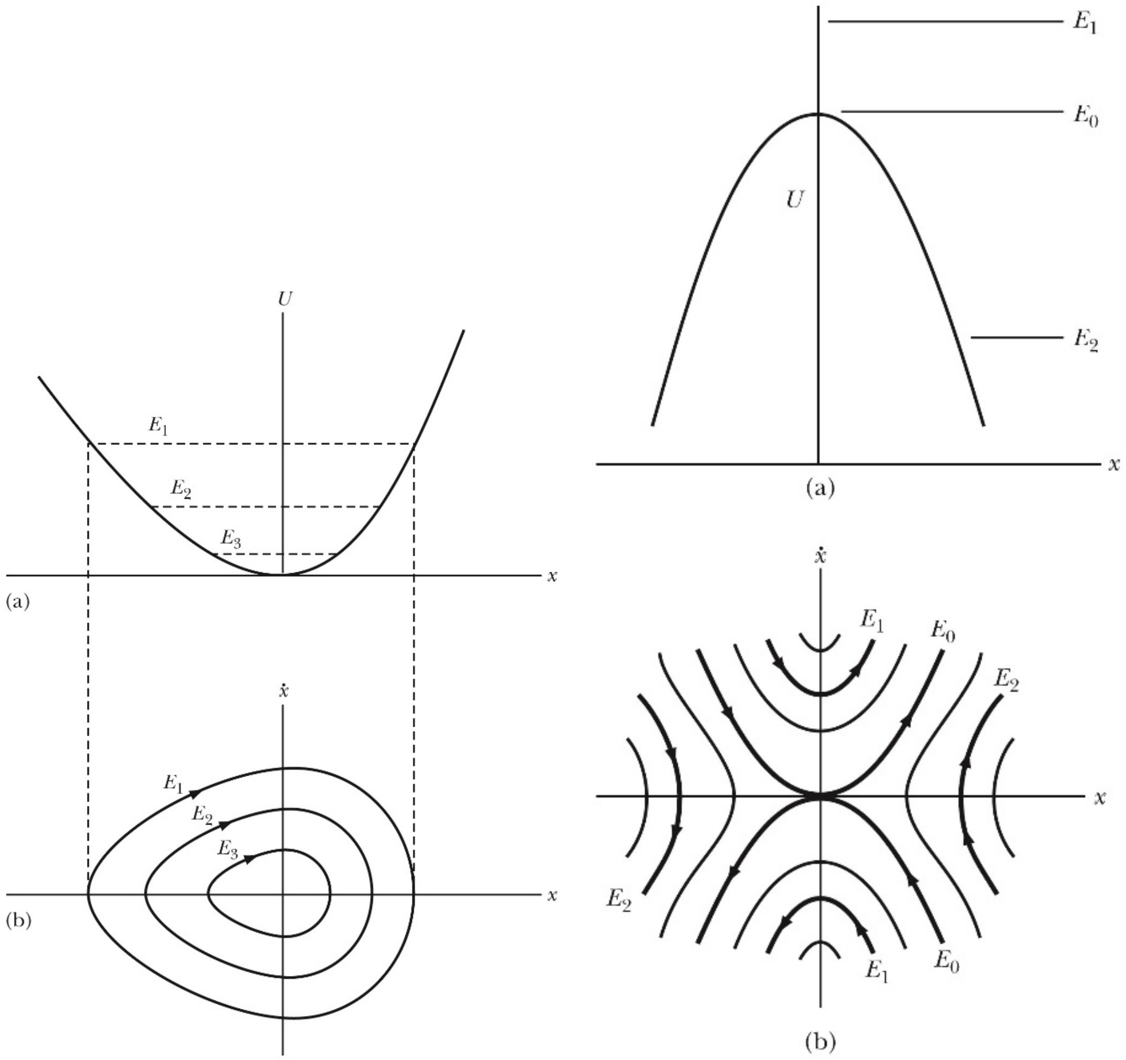


Figure 4.3: (a) Potential energy distributions around stable (left) and unstable (right) equilibrium positions. (b) Phase diagrams for stable (left) and unstable (right) equilibrium positions.

Examples of phase diagrams for stable and unstable equilibrium positions are shown in Fig. 4.3. The curves of constant energy are closed when we are dealing with a stable equilibrium, while they are not closed when we are dealing with unstable equilibrium.

4.3 The van der Pol Equation

An important type of non-linear differential equation was used to describe the non-linear oscillations in electronic circuits containing vacuum tubes. The van der Pol differential equation has the following form:

$$\ddot{x} + \mu\left(x^2 - a^2\right)\dot{x} + \omega_0{}^2 x = 0 \tag{4.6}$$

This equation is non-linear due to the presence of the x^2 term. One can interpret this equation as the equation of a system that has variable damping. In addition, the "damping" coefficient can be positive or negative, depending on the value of a. Based on what we have seen for damping in Chapter 3, we expect the following behavior for the solution of this differential equation:

1. If $|x| > |a|$, the “damping” coefficient is positive and the amplitude of the motion decreases.
2. If $|x| < |a|$, the “damping” coefficient is negative and the amplitude of the motion increases.

After a period of time we can imagine that the “damping” coefficient approaches zero and simple harmonic motion, with constant amplitude, will result. The curve mapped out in phase space in this case is called the **limit cycle**, and eventually no matter what the starting point, the system will follow the limit cycle.
To study the time evolution of the van der Pol equation, we need to use numerical methods to solve the differential equation. I will illustrate the use of Mathematica to solve this equation. The following code will allow you to play with the solutions of the van der Pol equation (the code can be found in the file *vanDerPolSolution* in the Mathematica folder under Computer Code on our website):

```
(* Set the values of the various parameters *)
mu = 0.05;
a = 1.0;
w0 = 1;

(* Solve the differential equations with the given set of initial \
conditions. *)
sol = NDSolve[
   {x'[t] == v[t],
    v'[t] == -mu*(x[t]*x[t] - a*a)*v[t] - w0*w0*x[t],
    x[0] == 1,
    v[0] == 0},
   {x, v}, {t, 0, 50*2*Pi}, MaxSteps -> 20000];

(* Plot the solution of the differential equations *)
ParametricPlot[Evaluate[{x[t], v[t]} /. sol], {t, 0, 50*2*Pi},
 PlotRange -> All,
 AxesLabel -> {"x (rad)", "v (rad/s)"}]
```

The results of this type of calculations, for different starting parameters and damping terms, are shown in Figs. 4.4 and 4.5. Fig. 4.4 shows that the limit cycle is the same for different initial values of x_0. Looking at Fig. 4.5, we observe that the shape of the limit cycle is a function of the damping parameter. Figure 4.6 shows that an increase of the damping parameter decreases the time required to approach the limit cycle.

4.4 The Plane Pendulum

A good example of a system with non-linear behavior is the plane pendulum. In many introductory physics courses, the simple pendulum is used to introduce harmonic motion. However, the simple pendulum will only exhibit simple harmonic motion if the angular displacement is small (small enough to ensure that $\sin\theta$ is approximately θ). If we can not make the assumption that the angular displacements are small, the simple pendulum will show non-linear behavior. The differential equation describing the system is given by

$$\ddot{\theta} + {\omega_0}^2 \sin\theta = 0 \tag{4.7}$$

where

$${\omega_0}^2 = \frac{g}{l} \tag{4.8}$$

and l is the length of the pendulum.
Before we look in detail at the phase diagrams for this pendulum, we should first consider what we expect to see. Let’s assume that at time $t = 0$ s, the pendulum is located at its equilibrium position, and it has a linear velocity v_0 at that time.

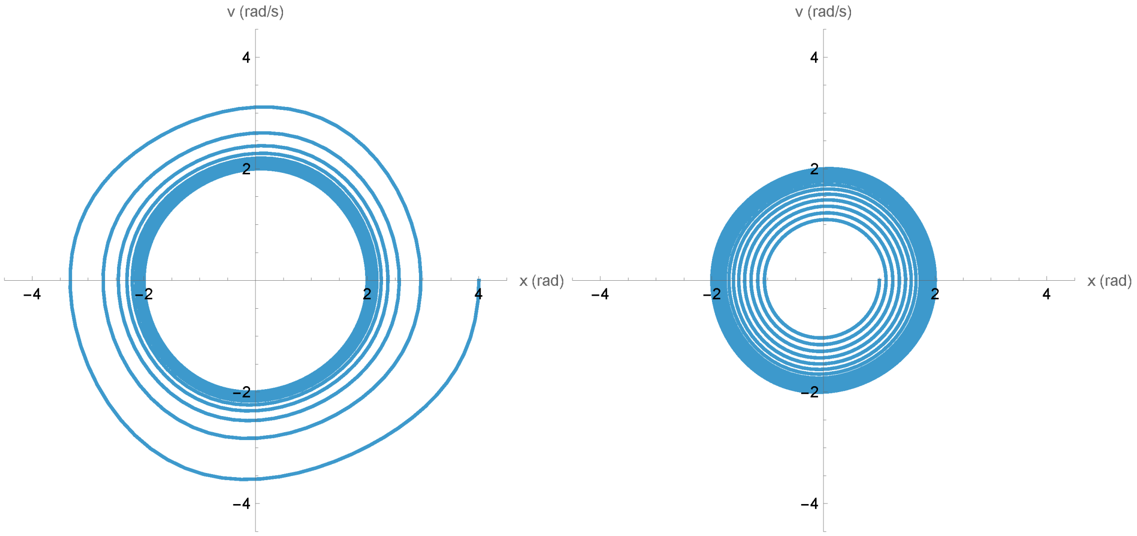


Figure 4.4: Solution of the van der Pol equation with $\mu = 0.05$, $a = 1$ rad, $\omega_0 = 1$ rad/s, $x_0 = 4$ rad, $v_0 = 0$ rad/s (left) and $\mu = 0.05$, $a = 1$ rad, $\omega_0 = 1$ rad/s, $x_0 = 1$ rad, $v_0 = 0$ rad/s (right).

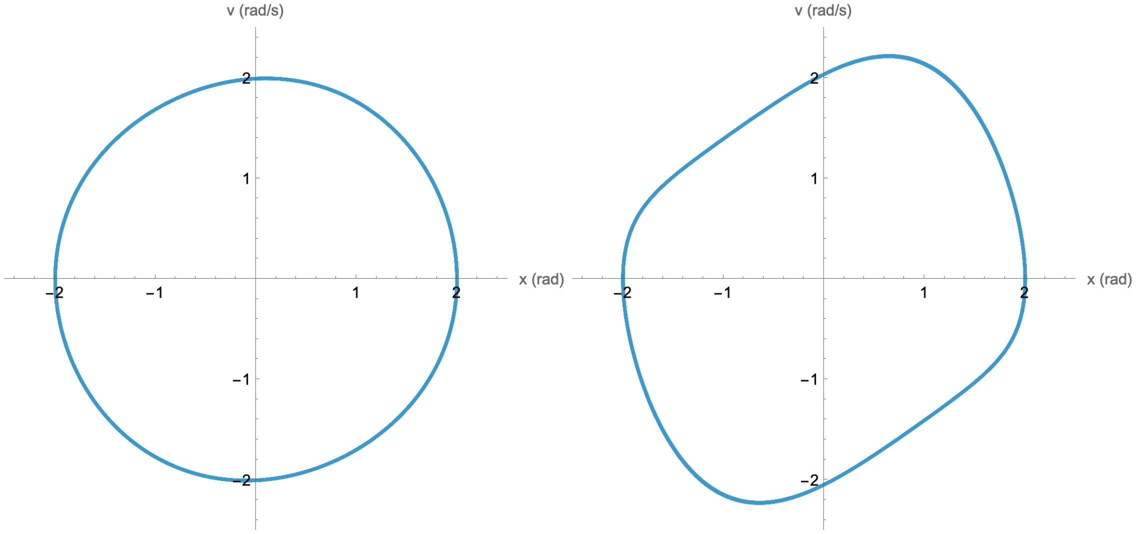


Figure 4.5: Solution of the van der Pol equation with $\mu = 0.05$, $a = 1$ rad, $\omega_0 = 1$ rad/s, $x_0 = 2$ rad, $v_0 = 0$ rad/s (left) and $\mu = 0.5$, $a = 1$ rad, $\omega_0 = 1$ rad/s, $x_0 = 2$ rad, $v_0 = 0$ rad/s (right).

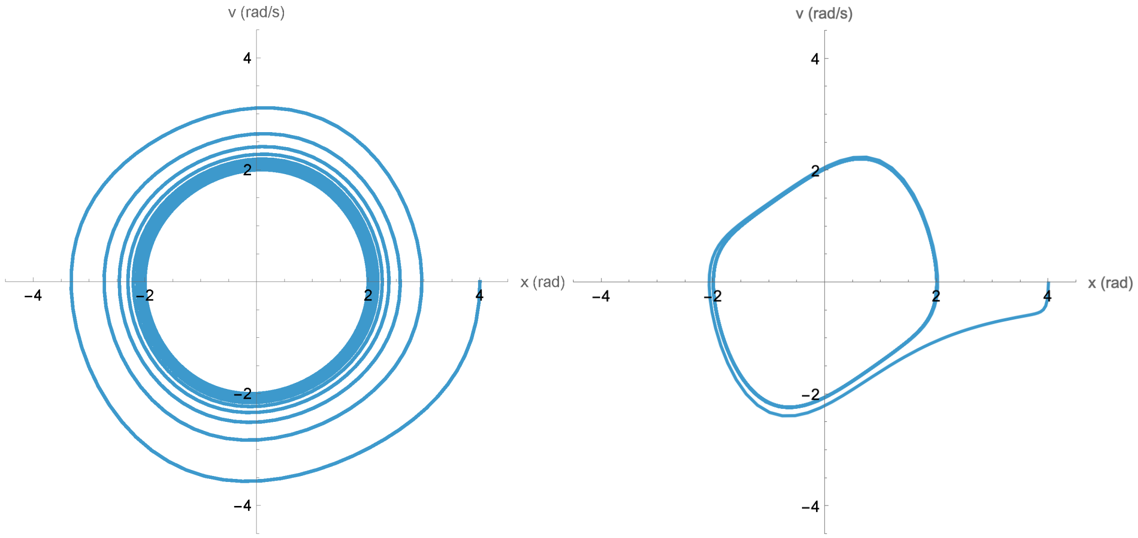


Figure 4.6: Solution of the van der Pol equation with $\mu = 0.05$, $a = 1$ rad, $\omega_0 = 1$ rad/s, $x_0 = 4$ rad, $v_0 = 0$ rad/s (left) and $\mu = 0.5$, $a = 1$ rad, $\omega_0 = 1$ rad/s, $x_0 = 4$ rad, $v_0 = 0$ rad/s (right).

1. If the velocity at $t = 0$ is small, the system will oscillate back and forth between a minimum and maximum value of the angular displacements, and the phase diagram is a closed curve (see Fig. 4.7).
2. If the velocity at $t = 0$ is sufficiently large, the system will continue to rotate in one direction, and the angular velocity will show an oscillatory behavior (but without a sign reversal). An example of this response is shown in Fig. 4.8.
3. The limiting case will be the case where the velocity at $t = 0$ is such that the pendulum makes it just to its highest possible position ($y = 2l$). Assuming the mass of the pendulum is connected via a weightless rod to the pivot point, the total energy of the mass will have to be $2mgl$ (kinetic energy will be equal to 0 at the top position). The kinetic energy at the equilibrium must therefore be $2mgl$, and the linear velocity at that postion will be equal to $2\sqrt{gl}$. The angular velocity at the equilibrium position will thus be equal to $(2\sqrt{gl}/l) = 2\sqrt{g/l} = 2\omega_0$. For the phase diagrams shown in Figs. 4.7 and 4.8, we have assumed $\omega_0 = 1$ rad/s, and the "critical" angular velocity is thus equal to 2 rad/s.

Figures 4.7 and 4.8 were made using Mathematica and can be reproduced using the following code, which is also contained in Mathematica notebook *planePendulum* that is available from our website:

```
(* Set the values of the various parameters *)
w0 = 1;

(* Solve the differential equations with the given set of initial \
conditions. *)
sol = NDSolve[
   {x'[t] == v[t],
    v'[t] == -w0*w0*Sin[x[t]],
    x[0] == 0,
    v[0] == 1.5},
   {x, v}, {t, 0, 50*2*Pi}, MaxSteps -> 20000];

(* Plot the solution of the differential equations *)
ParametricPlot[Evaluate[{x[t], v[t]} /. sol], {t, 0, 5*2*Pi},
 AxesLabel -> {"x (rad)", "v (rad/s)"}]
```

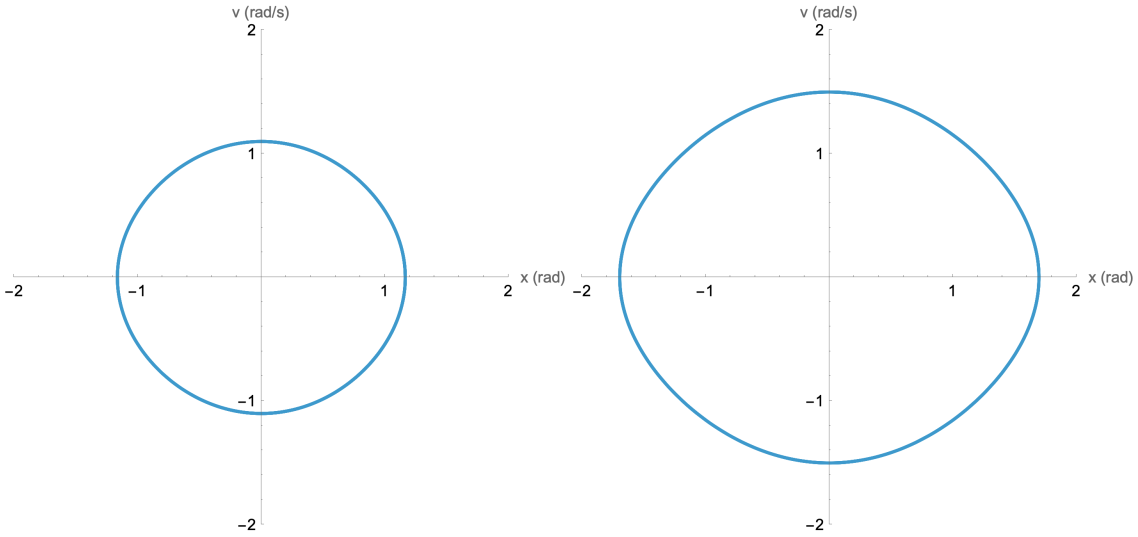


Figure 4.7: Solution of the differential equation associated with the plane pendulum with $\omega_0 = 1$ rad/s, $x_0 = 0$ rad, $v_0 = 1$ rad/s (left) and $\omega_0 = 1$ rad/s, $x_0 = 0$ rad, $v_0 = 1.5$ rad/s (right). **Note that the velocity v_0 is below the critical velocity.**

Another way to look at the phase diagram is to use contour plots. Rather than solving the differential equation for a specific set of initial parameters, and thus a specific energy, the contour plot provides an overview of the solution for a range of initial conditions. Consider a point on the phase diagram corresponding to an angular position θ and an angular velocity ω. The kinetic and potential energies at this location are:

$$T = \frac{1}{2}I\omega^2 = \frac{1}{2}ml^2\dot{\theta}^2 \tag{4.9}$$

and

$$U = mgl\,(1 - \cos\theta) \tag{4.10}$$

The total energy E is thus equal to

$$E = T + U = \frac{1}{2}ml^2\dot{\theta}^2 + mgl\,(1 - \cos\theta) \tag{4.11}$$

A contour plot, showing the total energy E as function of θ and ω, shows the nature of the motion that will occur when we give specific initial conditions for the pendulum. An example of a contour plot, obtained for $m = 1$ kg and $l = 1$ m is shown in Fig. 4.9. This contour plot was generated using the following Mathematica code, which is also contained in Mathematica notebook *planePendulumContours* that is available from our website:

```
l = 1;
g = 9.8;
ContourPlot[(1./2.)*l^2*w^2 + g*l*(1. - Cos[theta]),
 {theta, -4 Pi, 4 Pi}, {w, -15, 15},
 PlotLegends -> Automatic,
 FrameLabel -> {"theta (rad)", "omega (rad/s)"}]
```

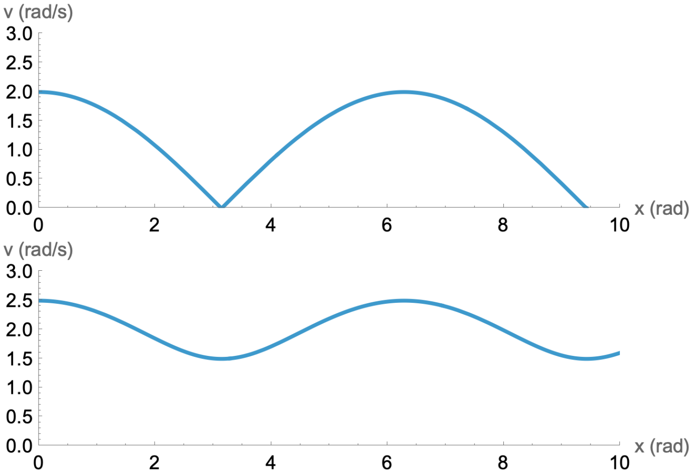


Figure 4.8: Solution of the differential equation associated with the plane pendulum with $\omega_0 = 1$ rad/s, $x_0 = 0$ rad, $v_0 = 2$ rad/s (left) and $\omega_0 = 1$ rad/s, $x_0 = 0$ rad, $v_0 = 2.5$ rad/s (right). **Note that the critical velocity v_0 is 2 rad/s.**

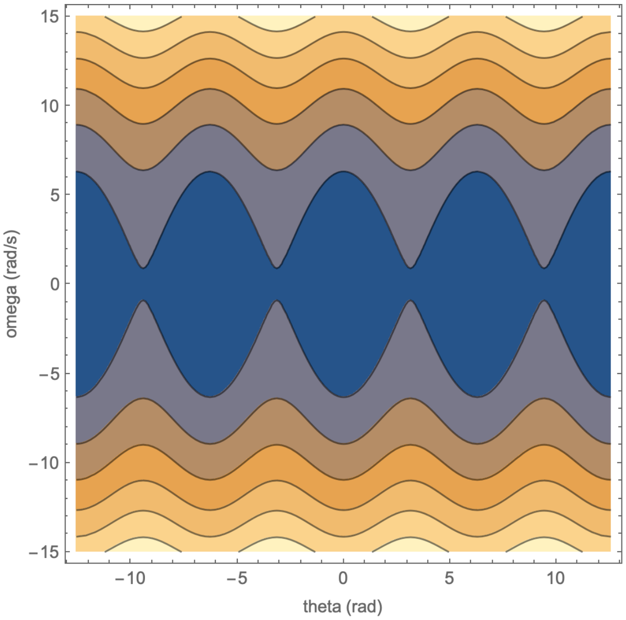


Figure 4.9: Contour plot for a plane pendulum ($m = 1$ kg, $l = 1$m).

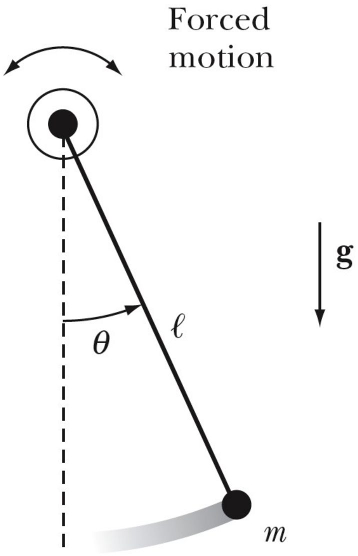


Figure 4.10: A damped pendulum, driven about its pivot point.

4.5 Chaos

Important principles of chaos can be studied using a simple driven pendulum. Consider the pendulum shown in Fig. 4.10, which is driven about its pivot point. We assume that the pendulum is damped with a damping coefficient b. The magnitude of the total torque around the pivot point is equal to

$$\tau = I\ddot{\theta} = ml^2\ddot{\theta} \tag{4.12}$$

The total torque is equal to the sum of the torques due to the gravitational force, the damping force, and the driving force:

$$\tau = -b\dot{\theta} - mgl\sin\theta + \tau_d\cos(\omega t) \tag{4.13}$$

The differential equation describing the motion of this driven pendulum is given by

$$\ddot{\theta} = -\frac{b}{ml^2}\dot{\theta} - \frac{g}{l}\sin\theta + \frac{\tau_d}{ml^2}\cos(\omega t) \tag{4.14}$$

This differential equation is a non-linear equation due to the $\sin\theta$ term. This equation has the following general form:

$$\ddot{x} = -c\dot{x} - \sin x + F\cos(\omega t) \tag{4.15}$$

The solution to this equation can be studied using numerical methods, and we will focus on the results obtained with Mathematica. The code required to study the solutions of this equation can be found in the file *chaosPendulum* in the Mathematica folder under Computer Code on our website:

```
(* Set the values of the various parameters *)
c = 0.2;
w = 0.694;
F = 0.52;
(* Increase F in steps of 0.1 to see the different solutions in the chapter notes. *)
pi = N[Pi];
cycles = 50;
```

```
steps = 30;

(* Solve the differential equations with the given set of initial \
conditions. *)
sol = NDSolve[
   {x'[t] == v[t],
    v'[t] == -c*v[t] - Sin[x[t]] + F*Cos[w*t],
    x[0] == 0.8,
    v[0] == 0.8},
   {x, v}, {t, 0, (cycles*(2 Pi)/w)}, MaxSteps -> 20000];

(* Create a function "reduce" that translate all angles back to the \
region between -Pi and +Pi *)
reduce[x_] := Mod[x, 2 pi] /; Mod[x, 2 pi] <= pi;
reduce[x_] := (Mod[x, 2 pi] - 2 pi) /; Mod[x, 2 pi] > pi;

(* Create a table with the angles mapped to the -Pi to Pi region. *)
points = Flatten[
   Table[
    {x[t], v[t]} /. sol,
    {t, 0, (cycles*(2 Pi)/w), (1/steps) (2 Pi/w) // N}
    ], 1
   ];
xposition = Table[{i, 1}, {i, Length[points]}];
newpoints = MapAt[reduce, points, xposition];

(* Plot the solution of the differential equations: v[t] vs t. *)
ParametricPlot[
 Evaluate[{t, v[t]} /. sol], {t, 0, (cycles*(2 Pi)/w)},
 PlotRange -> {{((cycles - 20)*(2 Pi)/w), (cycles*(2 Pi)/w)}, {-3, 3}},
 AxesLabel -> {"t/t0", "v (rad/s)"},
 AspectRatio -> 1/2
 ]

(* Plot the phase diagram. *)
ParametricPlot[
 Evaluate[{x[t], v[t]} /.
   sol], {t, ((cycles - 20)*(2 Pi)/w), (cycles*(2 Pi)/w)},
 AxesLabel -> {"x (rad)", "v (rad/s)"},
 AspectRatio -> 1/2
 ]

(* Plot the phase diagram (using the reduced angle). *)
ListPlot[
 newpoints,
 PlotRange -> {{-3, 3}, {-3, 3}},
 AxesLabel -> {"x (rad)", "v (rad/s)"},
 AspectRatio -> 1/2
 ]

(* Make the poincare plot. *)
poincare =
  Table[newpoints[[n]], {n, 1 + 25 steps, Length[points], steps}];
Length[poincare];
```

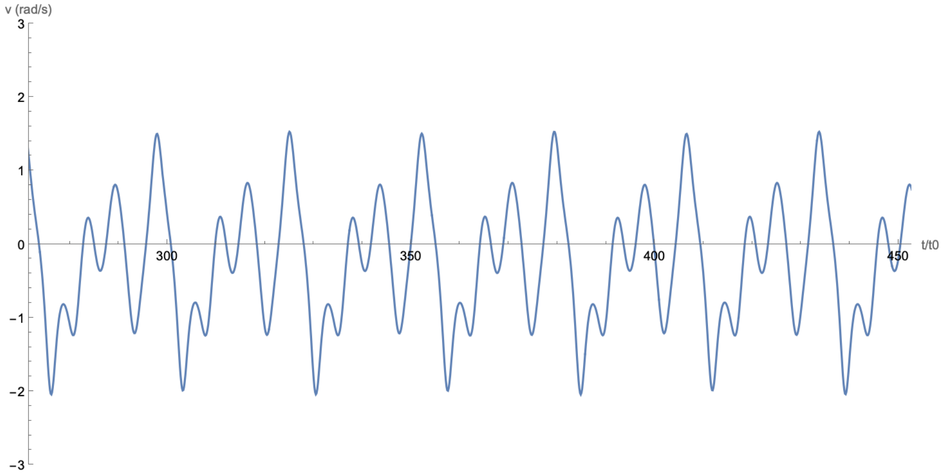


Figure 4.11: Velocity versus time for the driven pendulum with $c = 0.2$, $\omega = 0.694$, and $F = 0.52$.

```
ListPlot[
 poincare,
 PlotRange -> {{-3, 3}, {-3, 3}},
 AxesLabel -> {"x (rad)", "v (rad/s)"},
 PlotStyle -> PointSize[0.015],
 AspectRatio -> 1/2
 ]
```

There are many different ways of looking at the results of this calculation. Consider the case where $c = 0.2$, $\omega = 0.694$, and $F = 0.52$. A graph showing the velocity versus time is shown in Fig. 4.11. The period shown in Fig. 4.11 does not include the initial period when transient effects may dominate. The motion is periodic, but not simple harmonic. The average velocity is negative, indicating that on average the solution will move to more negative angles.

Another way to look at the solution is to construct the phase diagram. The phase diagram for the solution shown in Figure 11 is shown in Fig. 4.12. Figure 4.12(left) shows the velocity versus angle. As expected, the system will move towards more negative angles. Figure 4.12(right) shows the similar information to what is shown in Fig. 4.12(left), except that the angle is reduced to an angle between $-\pi$ and $+\pi$.

The phase diagrams shown in Fig. 4.12 are often difficult to interpret. An approach that was developed to emphasize the possible periodic nature of the solution is the Poincare representation. In this representation, we plot the velocity versus position at times separated by one period ($2\pi/\omega$. If the motion is periodic, each new point in the Poincare diagram will be plotted at the same position as the previous point: the plot shows thus a single point. **Note: the number of points shown will provide information about the periodicity; the actual position does not provide much information.** The Poincare representation of the phase diagrams shown in Fig. 4.12 is shown in Fig. 4.13. There are 3 points in the Poincare representation, suggesting a periodic solution (but not simple harmonic).

The solution of the driven pendulum is very sensitive to the parameters that are being used in the calculation. The transition of chaos is characterized by an increase in the number of points in the Poincare diagrams. Consider what happens if we change one of the parameters of our differential equation. Figure 4.14 shows the results of a calculation when we increase F from 0.52 to 0.62. The solution has become a simple harmonic, as can be seen by the fact that the Poincare plot contains only a single point.

Figures 4.15, 4.16, and 4.17 show the results of increasing the force F to 0.72, 0.82, and 0.92. The solution is chaotic when $F = 0.72$ and $F = 0.82$, but the solution is periodic when $F = 0.92$.

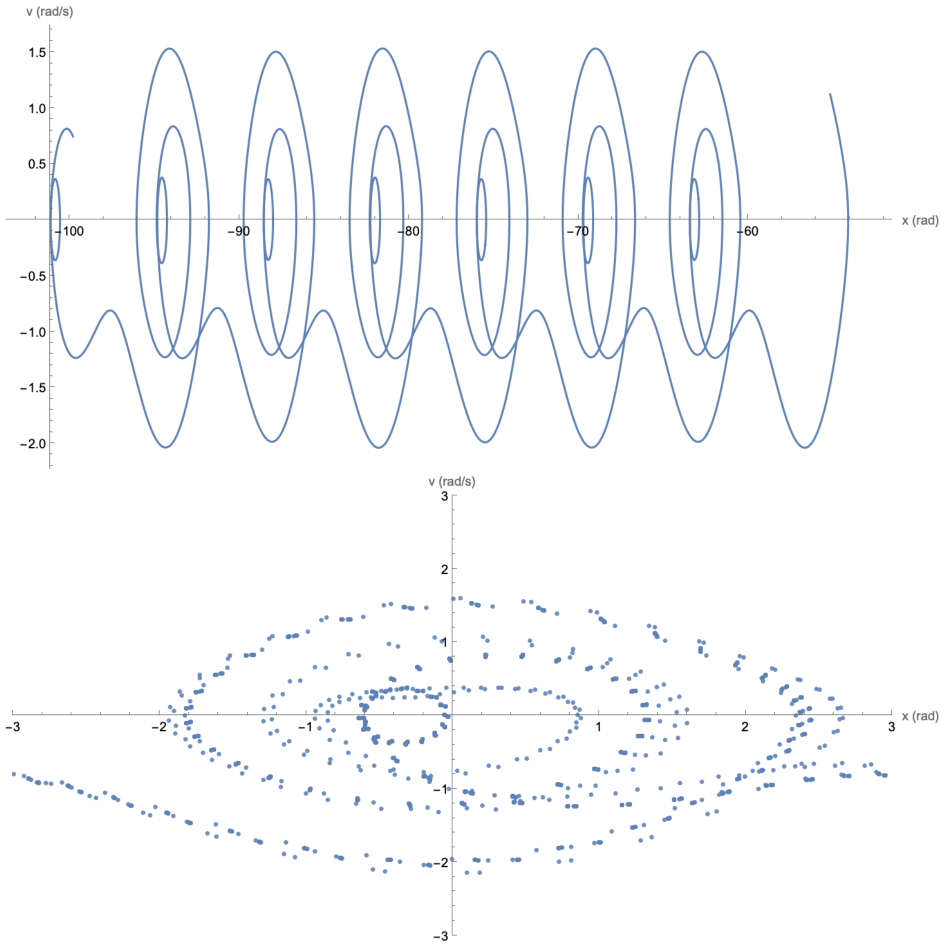


Figure 4.12: Phase diagrams for the driven pendulum with $c = 0.2$, $\omega = 0.694$, and $F = 0.52$ (left). The diagram shown on the right contains the same information as in a) except that the angles are mapped on the range between $-\pi$ and $+\pi$

.

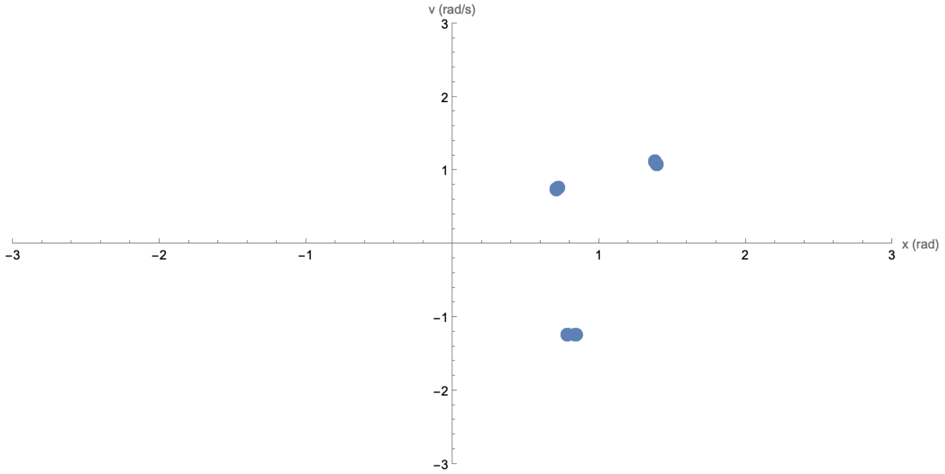


Figure 4.13: Poincare representation of the motion of a driven pendulum with $c = 0.2$, $\omega = 0.694$, and $F = 0.52$.

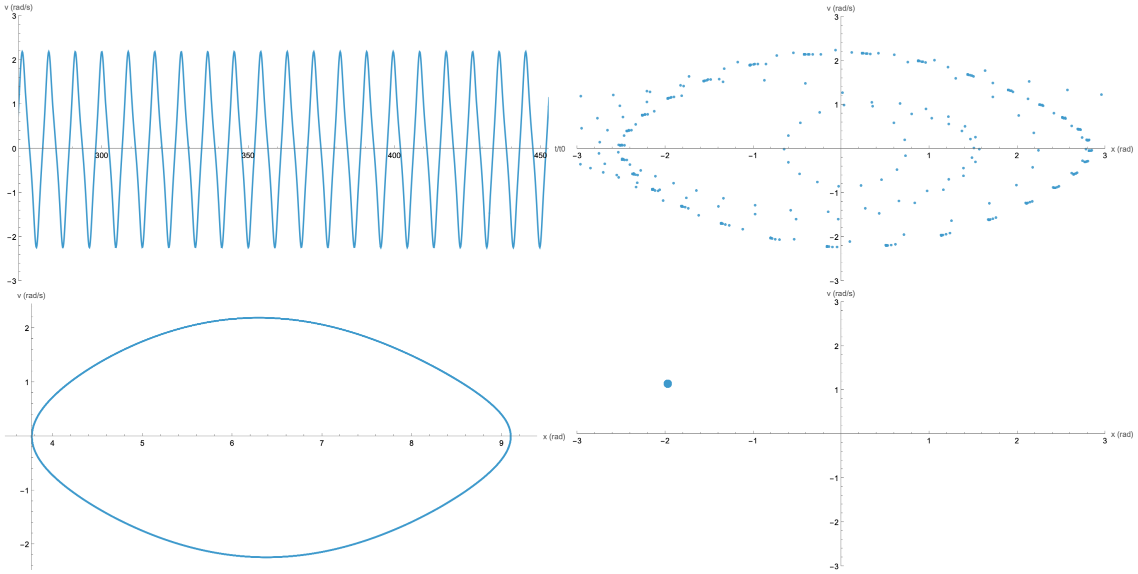


Figure 4.14: Solutions of the motion of a driven pendulum with $c = 0.2$, $\omega = 0.694$, and $F = 0.62$.

.

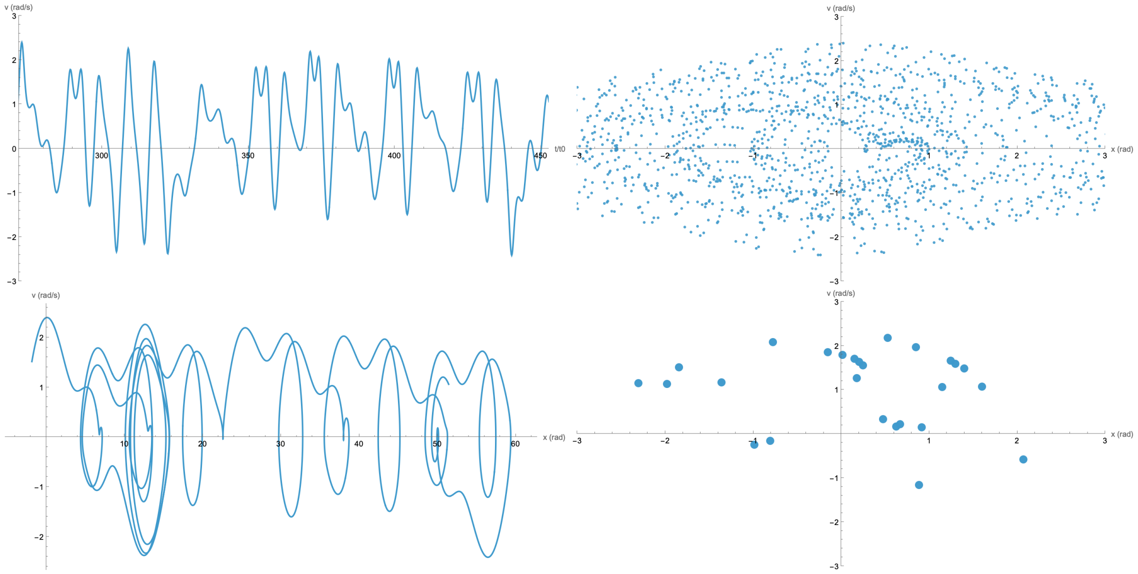


Figure 4.15: Solutions of the motion of a driven pendulum with $c = 0.2$, $\omega = 0.694$, and $F = 0.72$.

.

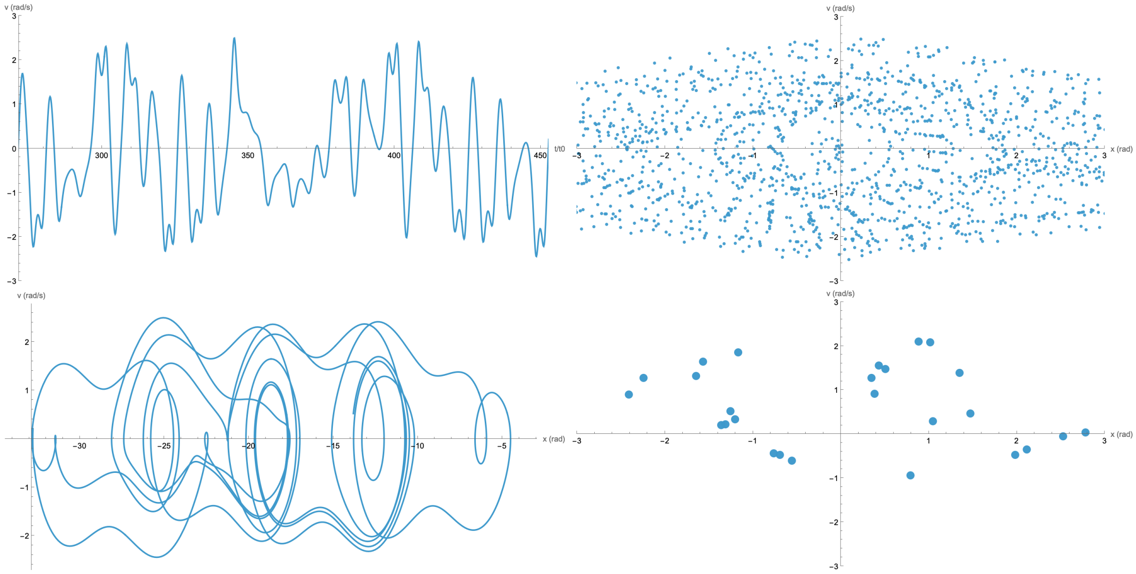


Figure 4.16: Solutions of the motion of a driven pendulum with $c = 0.2$, $\omega = 0.694$, and $F = 0.82$.

.

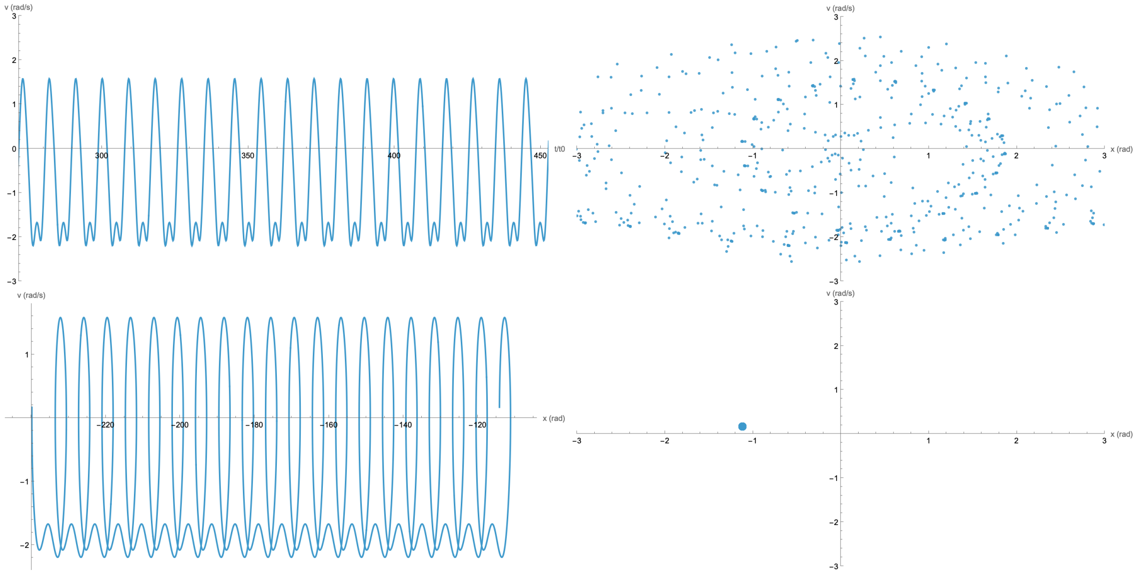


Figure 4.17: Solutions of the motion of a driven pendulum with $c = 0.2$, $\omega = 0.694$, and $F = 0.92$.

.

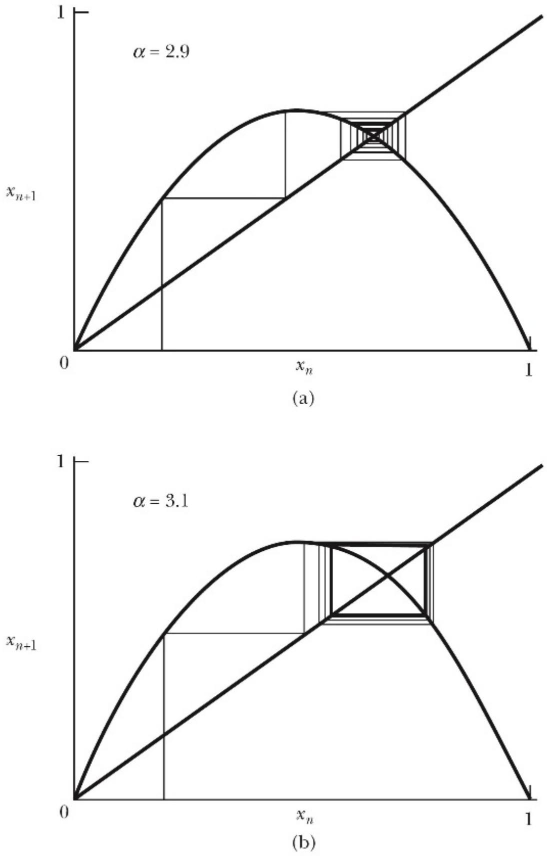


Figure 4.18: Examples of the logistic equation map for two different values of α.

4.6 Logistic Equations

Chaos can also develop when we work with maps. Maps are used to describe the evolution of a system, and can developed if we know how the future depends on the immediate past. For example, consider a situation where the value of a system parameter at time t_{N+1}, x_{N+1}, depends on the value of this parameter at time t_N, x_N:

$$x_{N+1} = \alpha x_n (1 - x_n) \tag{4.16}$$

The function describing this relation is called the **logistic equation**, and it depends on both x and on α:

$$f(\alpha, x) = \alpha x (1 - x) \tag{4.17}$$

For certain values of α, the system will evolve towards a stable equilibrium while for other starting value the system may evolve towards more than one possible solution. Two examples, for two different values of α, are shown in Figure 16.

In the example shown in Fig. 4.18b, there are two solutions to the logistic map. For higher value of α there may be more solutions. The best way to view the possible modes of evolution is to use a bifurcation diagram. In a bifurcation diagram, we plot the value of x_n as function of the parameter α. We can construct a bifurcation diagram using numerical methods, and we will focus on the results obtained with Mathematica. The code required to study the solutions of this equation can be found in the file *bifurcation* in the Mathematica folder under Computer Code on our website. A bifurcation diagram for the logistic equation discussed in this section is shown in Fig. 4.19. The two diagrams in Fig. 4.19 show the bifurcation diagrams that are created for a starting value $x_0 = 0.1$ (left) and $x_0 = 0.9$ (right). We see that for values of α below 3 there is only one solution. For values of α between 3 and 3.45, there are two possible solutions.

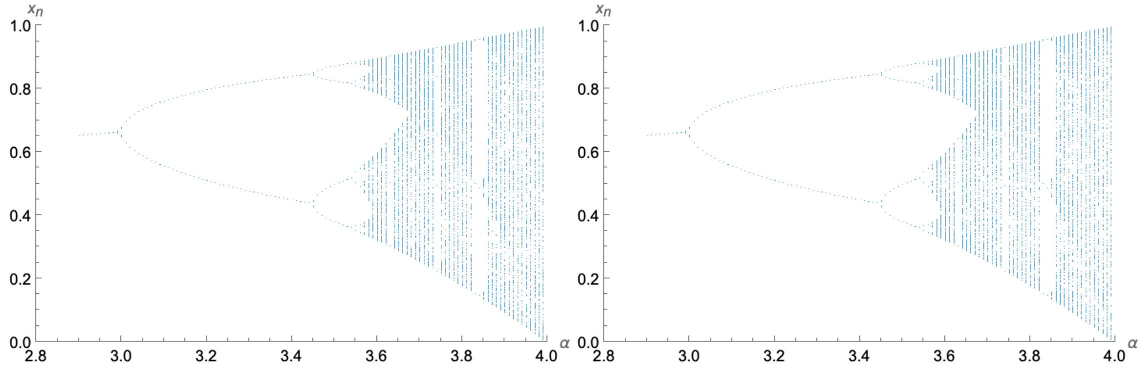


Figure 4.19: Bifurcation diagrams for $f(\alpha, x) = \alpha x(1-x)$ for a starting value of $x_0 = 0.1$ (left) and $x_0 = 0.9$ (right).

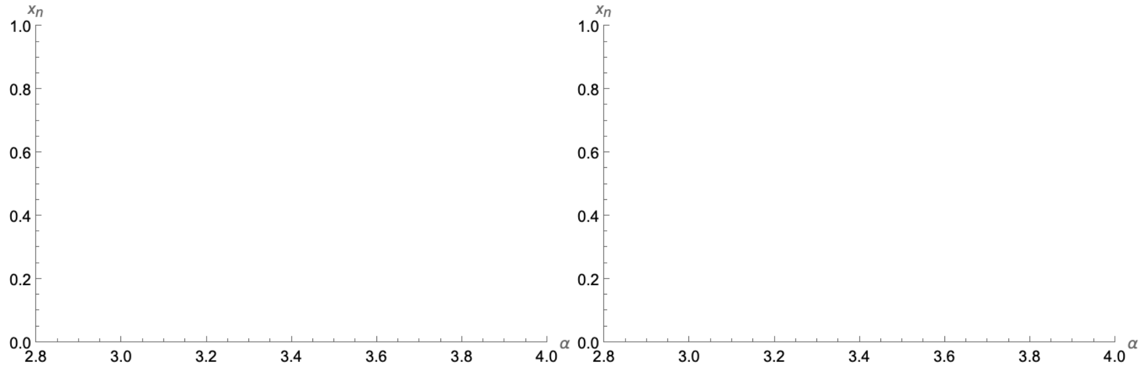


Figure 4.20: Bifurcation diagrams for $f(\alpha, x) = \alpha x(1-x)$ for a starting value of $x_0 = 0.0$) (left) and $x_0 = 1.0$) (right).

For values between 3.45 and 3.55 there are four possible solutions. Above $\alpha = 3.55$ the system has many possible solutions, and the system is said to be chaotic. However, as Fig. 4.19 shows, for very specific values of α the system returns back to a highly-ordered state with a small number of solutions. As can be seen, the bifurcation diagram does not change when we change the initial conditions, except for the special cases shown in Fig. 4.20 where $x_0 = 0.0$ (left) and $x_0 = 1.0$ (right).

Changes to the logistic equation will change the bifurcation diagram. Consider for example the following logistic equation:

$$f(\alpha, x) = \alpha x(1-x)^2 \tag{4.18}$$

The corresponding bifurcation diagram is shown in Fig. 4.21. Many of the features are similar, except that we see that the first doubling of the number of solutions does not occur until α exceeds 4.

When the time evolution of a system depends sensitively on the initial conditions, the system is said to show the characteristics of chaos. In the bifurcation diagrams shown in Fig. 4.19, we see that a particular system can have non-chaotic behavior for a specific range of parameters and chaotic behavior for a different range of parameters. A good way to examine the development of chaos is by looking at the Lyapunov exponent λ. The Lyapunov exponent is based on the differences between solutions if we make a small change in the initial conditions. Consider two starting values of x: $x_0 + \varepsilon$ and x_0. The difference between the solutions for these starting conditions is

$$d^n = f^n(\alpha, x_0 + \varepsilon) - f^n(\alpha, x_0) \tag{4.19}$$

It turns out that this difference grows as

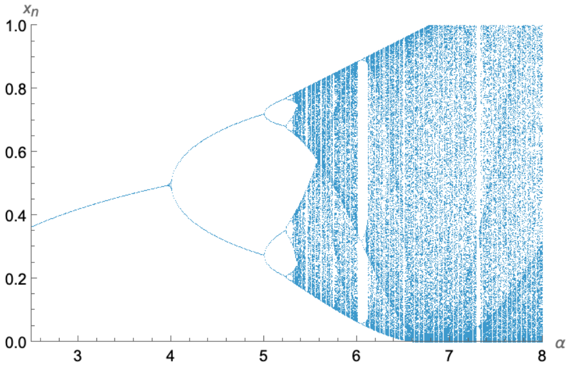


Figure 4.21: Bifurcation diagrams for $f(\alpha, x) = \alpha x(1-x)^2$ for a starting value of $x_0 = 0.1$.

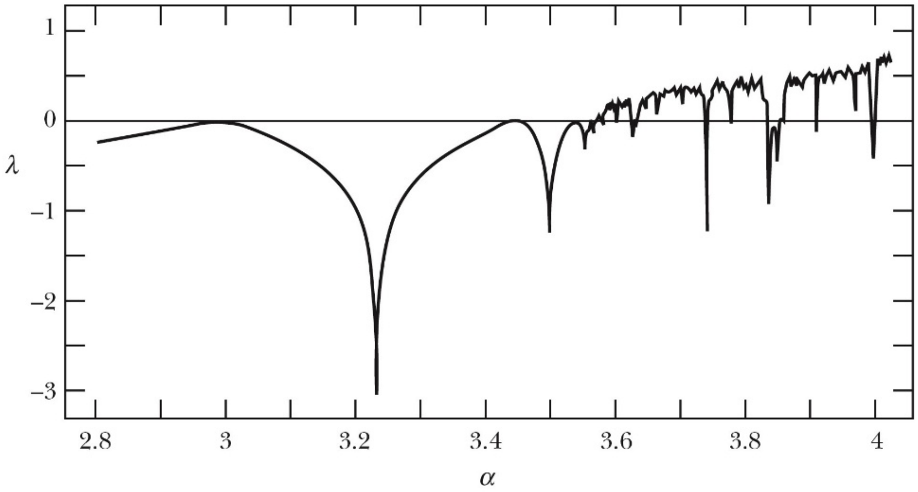


Figure 4.22: Lyapunov exponent as function of α for the logistic equation $f(\alpha, x) = \alpha x(1-x)$.

$$d^n = \varepsilon e^{n\lambda} \tag{4.20}$$

Combining these equations we see that

$$\lambda = \frac{1}{n}\ln\left(e^{n\lambda}\right) = \frac{1}{n}\ln\left[\frac{f^n(\alpha, x_0+\varepsilon) - f^n(\alpha, x_0)}{\varepsilon}\right] = \frac{1}{n}\ln\left[\left.\frac{df^n(\alpha,x)}{dx}\right|_{x_0}\right] \tag{4.21}$$

Using the chain rule for derivations, we can rewrite this equation as

$$\lambda = \frac{1}{n}\ln\left[\left.\frac{df^n(\alpha,x)}{dx}\right|_{x_0}\right] = \frac{1}{n}\ln\left[\left.\frac{df(\alpha,x)}{dx}\right|_{x_{n-1}}\left.\frac{df(\alpha,x)}{dx}\right|_{xx_{n-2}}....\left.\frac{df(\alpha,x)}{dx}\right|_{x_1}\left.\frac{df(\alpha,x)}{dx}\right|_{x_0}\right] \tag{4.22}$$

If we take the limit of n to infinity, we get the definition of the Lyapunov exponent

$$\lambda = \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\ln\left|\frac{df(\alpha,x_i)}{dx}\right| \tag{4.23}$$

A plot of the Lyapunov exponent as function of α for the logistic function $f(\alpha, x) = \alpha x(1-x)$ is shown in Fig. 4.22.

If the Lyapunov exponent is negative, the solutions will eventually converge. In Fig. 4.22 we can easily identify the region where a stable solution is achieved. Bifurcation (doubling of the number of solutions) will occur when the Lyapunov exponent is zero. Chaos will occur when the Lyapunov exponent is larger than 0.

5 Gravitation

In this Chapter, we will review the properties of the gravitational force. The gravitational force has been discussed in great detail in your introductory physics courses, and we will primarily focus on specifying the properties of the force and its associated force field, using the vector notations we have introduced in the previous Chapters.

5.1 The Gravitational Force

The gravitational force between two point-like particles is proportional to the product of their masses, and inversely proportional to the square of the distance between them. The force is **always** attractive, and directed along the line connecting the two particles.

$$\vec{F} = -G\frac{mM}{r^2}\hat{r} \tag{5.1}$$

The constant G is the gravitational constant, whose value is 6.673 x 10^{-11} Nm^2/kg^2. This relation is known as Newton's law of universal gravitation. The principle of superposition allows us to calculate the force on mass m due to multiple other point-like masses $M_1, M_2, \ldots\ldots$:

$$\vec{F} = -Gm\sum_{i=1}^{n}\frac{M_i}{{r_i}^2}\hat{r}_i \tag{5.2}$$

where r_i is the distance between mass M_i and mass m. Equation 5.2 is correct only if both masses are point-like objects. If one of the masses has a continuous mass distribution, we must replace the sum with an integral:

$$\vec{F} = -Gm\int_V \frac{\rho(\vec{r}')\hat{r}'}{r'^2}dv' \tag{5.3}$$

5.2 The Gravitational Field

The gravitational field generated by the gravitational force is defined in much the same way as we defined the electrostatic field generated by the electrostatic force:

$$\vec{g} = \frac{\vec{F}}{m} = -G\int_V \frac{\rho(\vec{r}')\hat{r}'}{r'^2}dv' \tag{5.4}$$

The units of the gravitational field are N/kg = m/s^2.

5.3 The Gravitational Potential

The gravitational potential is defined as the scalar function Φ whose gradient is equal to the opposite of the gravitational field:

$$\vec{g} = -\vec{\nabla}\Phi \tag{5.5}$$

One way to determine if we can find such a scalar function is to calculate the curl of the gravitational field:

$$\vec{\nabla}\times\vec{g} = -\vec{\nabla}\times\vec{\nabla}\Phi = -\begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial\Phi}{\partial x} & \frac{\partial\Phi}{\partial y} & \frac{\partial\Phi}{\partial z} \end{vmatrix} = 0 \tag{5.6}$$

In the case of the gravitational field we find

$$\vec{\nabla}\times\vec{g} = -GM\left(\vec{\nabla}\times\frac{\hat{r}}{r^2}\right) = -GM\left\{\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\left(\frac{1}{r^2}\right)\widehat{\theta} - \frac{1}{r}\frac{\partial}{\partial\theta}\left(\frac{1}{r^2}\right)\widehat{\phi}\right\} = 0 \tag{5.7}$$

and we conclude that there is a scalar function that can generate the gravitational field. Since the gravitational field depends just on r, we expect that the gravitational potential is also just a function of r. The following scalar function can generate the gravitational field:

$$\Phi = -G\frac{M}{r} \tag{5.8}$$

This is the gravitational potential due to point mass M. If we have a continuous mass distribution, the gravitational potential will be equal to

$$\Phi = -G\int_V \frac{\rho\left(\vec{r}'\right)}{r'}dv' \tag{5.9}$$

In this equation, we have assumed that the constant of integration is equal to 0 (or that the gravitational potential is 0 at infinity).

One of the reasons that the gravitational potential is introduced is that even for extended mass distributions, it is in general easier to calculate the gravitational potential (which is a scalar) instead of the gravitational force (which is a vector). Once we have determined the gravitational potential, we can determine the gravitational force by calculating the gradient of the gravitational potential.

5.4 The Gravitational Potential Energy

The gravitational potential can be used to determine the gravitational potential energy U of a particle of mass m. Recall from Physics 121 or Physics 141 that the change in the potential energy of an object, when it moves from one position to another position, is the opposite of the work done by the forces acting on the object. Consider an object that moves in the gravitational field of a point mass M located at the origin of a coordinate system. The work done on the object of unit mass by the gravitational force is equal to

$$dW = -\vec{g}\cdot d\vec{r} = \left(\vec{\nabla}\Phi\right)\cdot d\vec{r} = \sum_{i=1}^{3}\frac{\partial\Phi}{\partial x_i}dx_i = d\Phi \tag{5.10}$$

If we move the object of unit mass from infinity to a specific position, the work done is equal to

$$W\left(\vec{r}\right) = \int_{\infty}^{\vec{r}} d\Phi = \Phi\left(\vec{r}\right) \tag{5.11}$$

The work done to move an object of mass m to this position is thus equal to

$$W\left(\vec{r}\right) = m\Phi\left(\vec{r}\right) \tag{5.12}$$

The potential energy of the object at this position is thus equal to

$$U\left(\vec{r}\right) = m\Phi\left(\vec{r}\right) \tag{5.13}$$

Information about the force on the object can be obtained by taking the gradient of the potential energy:

$$F\left(\vec{r}\right) = -\vec{\nabla}U\left(\vec{r}\right) = -\frac{\partial}{\partial r}\left(-G\frac{mM}{r}\right)\hat{r} = -G\frac{mM}{r^2}\hat{r} \tag{5.14}$$

which is of course equal to the gravitational force we used as our starting point.

5.5 Visualization of the Gravitational Potential

We can visualize the gravitational potential in a number of different ways. The most common ways are contour plots, showing equipotential surfaces, and 3D plots showing the gravitational potential as function of the two-dimensional position (x, y). A simple program that can be used to generate such plots is *gravitationalPotential* which can be found in the Mathematica folder under Computer Code on our website:

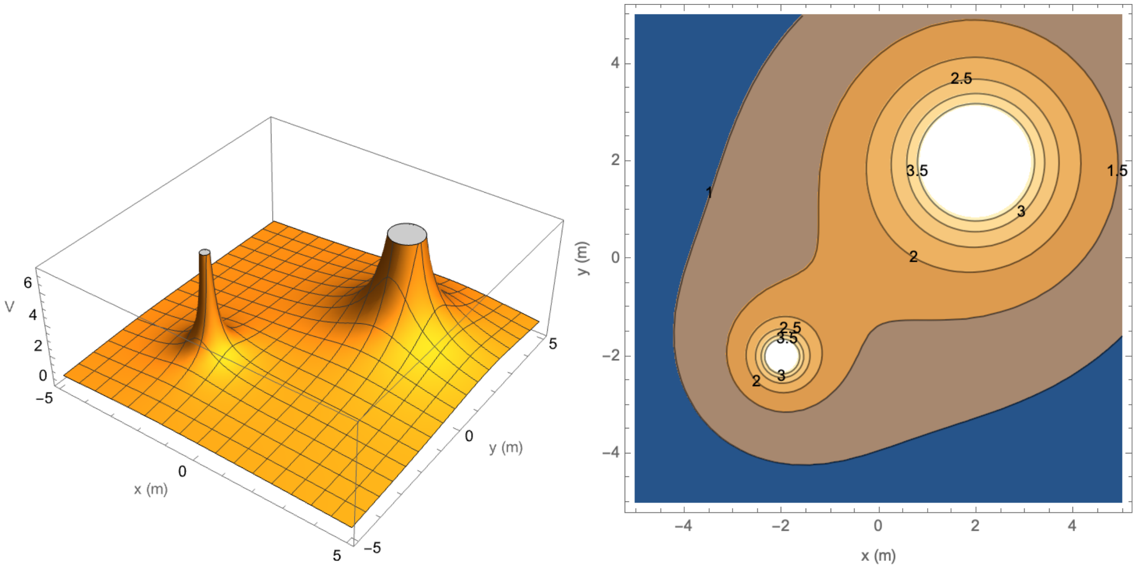


Figure 5.1: Different ways to visualize the gravitational potential due to two point masses of mass 1 and mass 3, located at (-2,-2) and (2,2), respectively.

```
(* Make a 3D plot of the gravitational potential due to two point \
masses of mass 1 and mass 4, located at (-2,-2) and (2,2), \
respectively. *)

Plot3D[(1/Sqrt[(x + 2)^2 + (y + 2)^2]) + (4/
    Sqrt[(x - 2)^2 + (y - 2)^2]),
 {x, -5, 5}, {y, -5, 5},
 AxesLabel -> {"x (m)", "y (m)", "V"},
 PlotRange -> {0, 7.5}, PlotPoints -> 50,
 Ticks -> {Automatic, Automatic, Automatic}]

(* Make a 2D contour plot of the gravitational potential due to two \
point masses of mass 1 and mass 4, located at (-2,-2) and (2,2), \
respectively. *)
ContourPlot[(1/Sqrt[(x + 2)^2 + (y + 2)^2]) + (4/
    Sqrt[(x - 2)^2 + (y - 2)^2]), {x, -5, 5}, {y, -5, 5},
 FrameLabel -> {"x (m)", "y (m)"}, ContourLabels -> True]
```

The plots that are generated using this program are shown in Fig. 5.1.

5.6 The Shell Theorem

When we calculate the gravitational force or the gravitational potential generated by a mass distribution, we can always use the most general expression for these quantities in terms of the volume integral over the mass distribution. However, if the mass distribution has spherical symmetry, we can use the shell theorem to calculate the gravitation force and the gravitational potential. The shell theorem states that **the gravitational potential at any point outside a spherically symmetric mass distribution is independent of the size of the distribution, and we can consider all of its mass to be concentrated at the center of the mass distribution. The gravitational potential is constant at any point inside a spherically symmetric mass distribution.** Using the shell theorem it is easy to calculate the gravitational potential and the gravitational

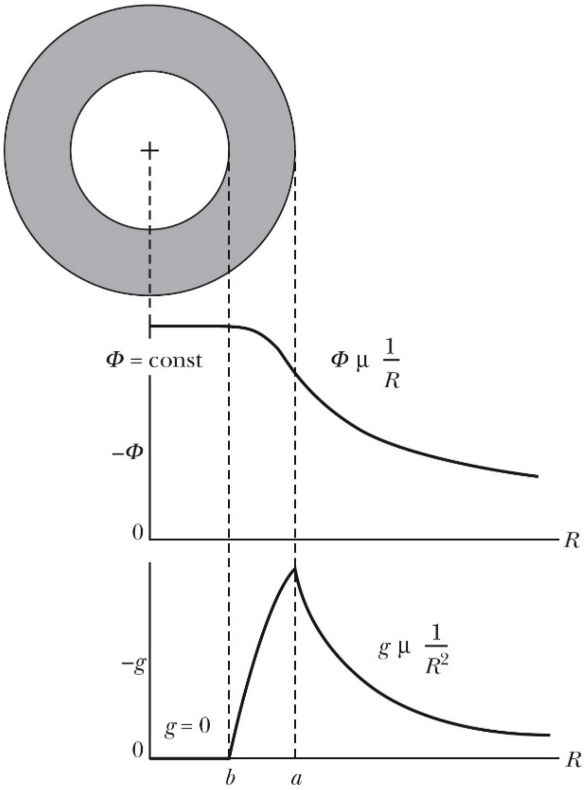


Figure 5.2: The gravitational potential and the gravitational force due to a spherical shell.

force due to a spherical shell, which is shown in Fig. 5.2.

The shell theorem can be used to make important predictions about orbital motion of planets around the central star and solar systems around the center of a galaxy. The observation of the rotational motion of the planets around the Sun yielded the first determination of its mass. Since the Sun is much more massive than any of the planets in the solar system, the motion of the planets could be described in terms of the gravitational force due to just the mass of the Sun (see Fig. 5.3).

The solar systems in most galaxies carry out an orbit around the center of the galaxy. Since it is assumed that a massive black hole is located in the center of most galaxies, we expect to see a trend in the orbital velocity versus distance, similar to the trend seen in our solar system (see Fig. 5.3). In reality, we see a distribution that decreases at a much smaller rate as function of the distance from the center of the galaxy, as shown in Fig. 5.4). This implies that there is more mass in the galaxy than was assumed, but also that this extra mass is **NOT** located in the center of the galaxy, but throughout the galaxy. This extra matter, that we can not see directly, is called dark matter and your teacher is looking for it in a gold mine in South Dakota.

5.7 The Poisson Equation

In Electricity and Magnetism we greatly benefited from the Gauss' law, which related the electric flux through a closed surface with the total charge enclosed by that surface. Given the fact that the nature of the gravitational force (its $1/r^2$ dependence) is similar to the nature of the electric force, we expect that similar "laws" apply to the gravitational force.

Based on the approach we took in Electricity and Magnetism, we define the gravitational flux due a point mass m in the following manner:

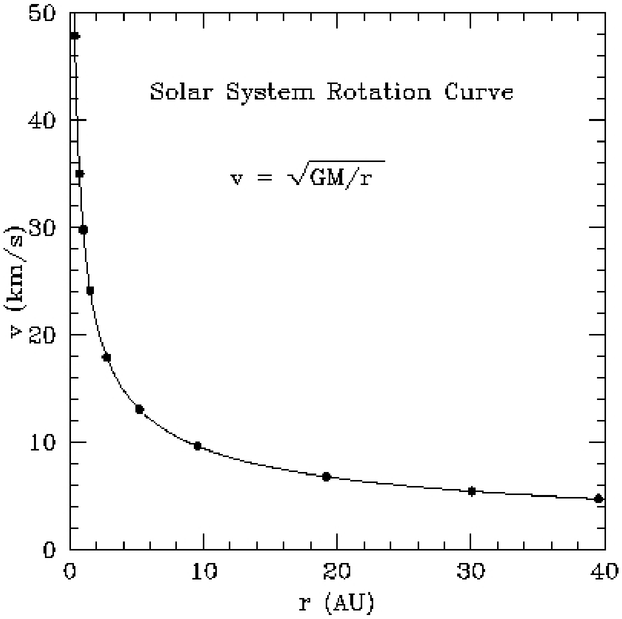


Figure 5.3: The orbital velocity of the planets in our solar system as function of their distance from the sun. The theoretical dependence, which depends only on the mass of the sun, is shown by the solid curve, and does an excellent job describing the trend in the data.

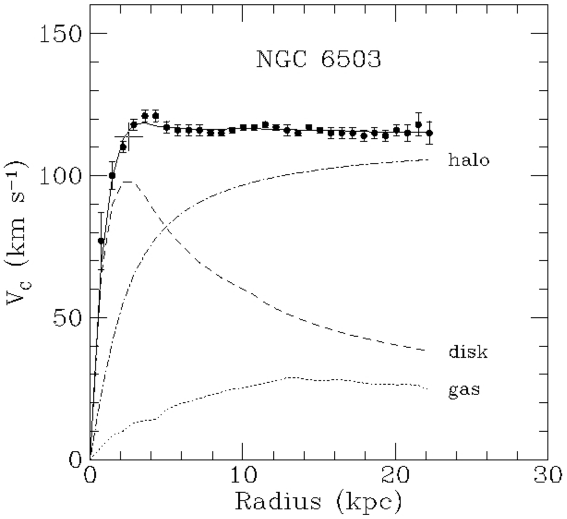


Figure 5.4: Measured orbital velocity of stars as function of distance from the center of the galaxy. The rotational curve can only be explained if we assume that there is halo of "dark matter", distributed throughout this galaxy.

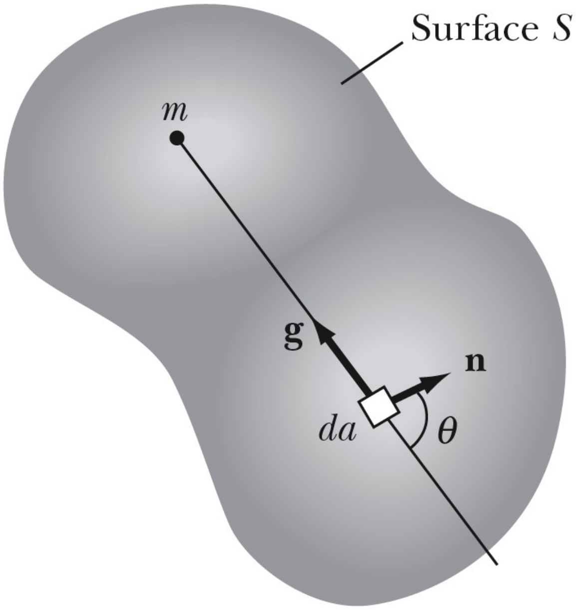


Figure 5.5: Surface used to calculate gravitational flux associated with a point mass m.

$$\Phi_{grav} = \int_S (\hat{n} \cdot \vec{g})\, da \tag{5.15}$$

where S is an arbitrary surface surrounding the mass m (see Fig.5.5). Using the definition of the gravitational field, we can rewrite this equation as

$$\Phi_{grav} = -Gm \int_S \frac{\cos\theta}{r^2} da = -Gm \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \left(\frac{\cos\theta}{r^2} \right) r^2 \sin\theta d\theta d\phi = -2\pi Gm \int_{\theta=0}^{\pi} \cos\theta \sin\theta d\theta = -4\pi Gm \tag{5.16}$$

When we have a mass distribution inside the surface S we need to replace m by a volume integral over the mass distribution:

$$\Phi_{grav} = -4\pi G \int_V \rho dv \tag{5.17}$$

The left-hand side of this equation can be rewritten using Gauss's divergence theorem:

$$\Phi_{grav} = \int_S (\hat{n} \cdot \vec{g})\, da = \int_V \left(\vec{\nabla} \cdot \vec{g} \right) dv \tag{5.18}$$

We thus conclude that

$$\int_V \left(\vec{\nabla} \cdot \vec{g} \right) dv + 4\pi G \int_V \rho dv = \int_V \left\{ \left(\vec{\nabla} \cdot \vec{g} \right) + 4\pi G\rho \right\} dv = 0 \tag{5.19}$$

for any volume V and thus

$$\left(\vec{\nabla} \cdot \vec{g} \right) = -4\pi G\rho \tag{5.20}$$

This equation can also be expressed in terms of the gravitational potential and becomes

$$\vec{\nabla} \cdot \left(-\vec{\nabla}\Phi \right) = -\vec{\nabla}^2 \Phi = -4\pi G\rho \tag{5.21}$$

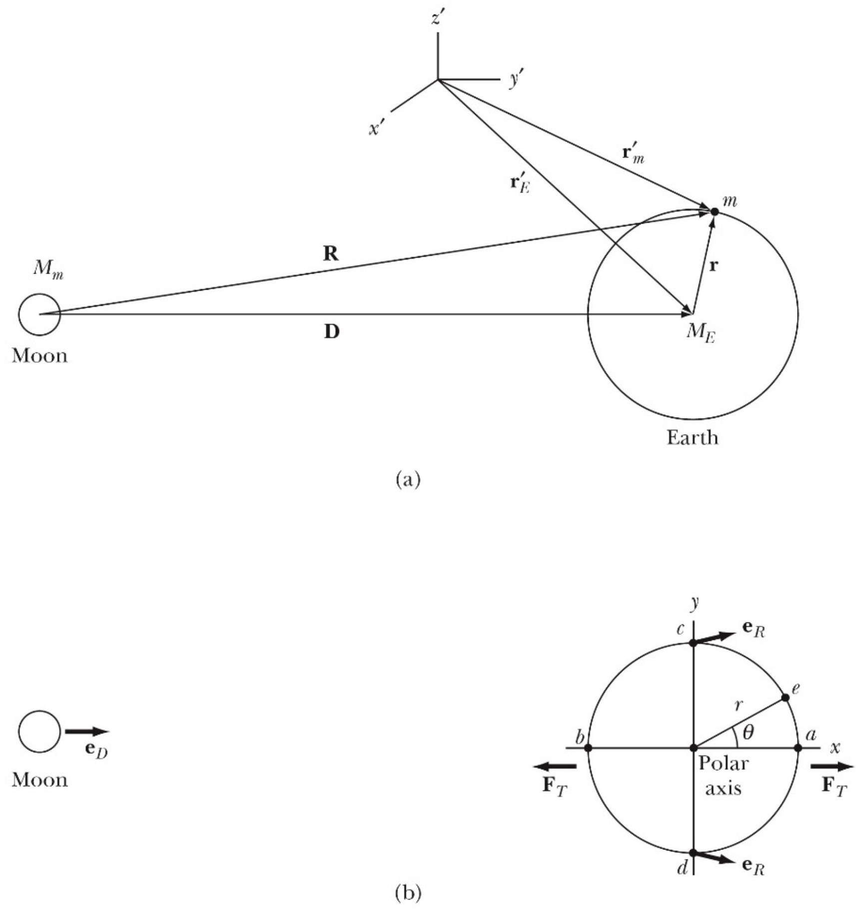


Figure 5.6: Geometry used to determine the forces on a volume of water of mass m, located on the surface of the Earth.

or

$$\vec{\nabla}^2 \Phi = 4\pi G \rho \tag{5.22}$$

This equation is known as **Poisson's equation**. The operator $\vec{\nabla}^2$ is called **the Laplacian**.

5.8 The Tides

We all know that the tides are caused by the motion of the moon around the Earth, but most of us have a more difficult time to explain why we have two high tides a day. Our simple picture would suggest that you have high tide on that side of the Earth closest to the moon, but this would only explain one high tide a day.

The calculation of the effect of the moon on the water on Earth is complicated by the fact that the Earth is not a good inertial system, so we have to assume that we can find a good inertial frame in which we can describe the forces acting on the water (see Fig. 5.6). Consider the forces on a volume of water of mass m, located on the surface of the Earth (see Fig. 5.6), due to the moon and the Earth:

$$\vec{F}_m = m\ddot{\vec{r}}_m = -\frac{GmM_E}{r^2}\hat{r} - \frac{GmM_m}{R^2}\hat{R} \tag{5.23}$$

The force exerted by the moon on the center of the Earth is equal to

$$\vec{F}_E = -\frac{GM_m M_E}{D^2}\hat{D} = -M_E \ddot{\vec{r}}_E \tag{5.24}$$

When we view the motion of the water on Earth, we view its motion with respect to the Earth. We have used a correct inertial frame to determine the acceleration of the individual components of our system, and we can now transform to a reference frame in which the Earth is at rest and centered around the origin (note: this is a non-inertial reference frame). In this reference frame, the acceleration of mass m is equal to

$$\ddot{\vec{r}} = \ddot{\vec{r}}_m - \ddot{\vec{r}}_E = -\frac{GM_E}{r^2}\hat{r} - \frac{GM_m}{R^2}\hat{R} + \frac{GM_m}{D^2}\hat{D} = -\frac{GM_E}{r^2}\hat{r} - GM_m\left(\frac{\hat{R}}{R^2} - \frac{\hat{D}}{D^2}\right) \tag{5.25}$$

The first term on the right-hand side is just the gravitational force on mass m due to the Earth. It will be the same anywhere on the surface of the Earth and is **not** responsible for the tides. The second part of the right-hand side is equal to the acceleration associated with the tidal force. It is related to the difference between the gravitational pull of the moon on the center of the Earth and on the surface of the Earth. Let us consider what this equation tells us about the magnitude and the direction of the tidal force at 4 different positions on the surface of the Earth (see Fig. 5.6).

1. Point a: At point a, the unit vectors associated with R and D are pointing in the same direction. Since $R > D$, the $(1/R^2)$ terms will be smaller than the $(1/D^2)$ term, and the tidal acceleration will be directed towards the right:

$$\ddot{\vec{r}}_T = GM_m\left(\frac{1}{D^2} - \frac{1}{R^2}\right)\hat{D} = GM_m\left(\frac{1}{D^2} - \frac{1}{(D+r)^2}\right)\hat{D} = \frac{GM_m}{D^2}\left(1 - \frac{1}{\left(1+\frac{r}{D}\right)^2}\right)\hat{D} \simeq 2GM_m\frac{r}{D^3}\hat{D} \tag{5.26}$$

2. Point b: At point b, the unit vector associated with R and D are pointing in the same direction. But, since $R < D$, the $(1/R^2)$ terms will be larger than the $(1/D^2)$ term, and the tidal acceleration will be directed towards the left but with the same magnitude as the acceleration at point a:

$$\ddot{\vec{r}}_T = GM_m\left(\frac{1}{D^2} - \frac{1}{R^2}\right)\hat{D} = GM_m\left(\frac{1}{D^2} - \frac{1}{(D-r)^2}\right)\hat{D} = \frac{GM_m}{D^2}\left(1 - \frac{1}{\left(1-\frac{r}{D}\right)^2}\right)\hat{D} \simeq -2GM_m\frac{r}{D^3}\hat{D} \tag{5.27}$$

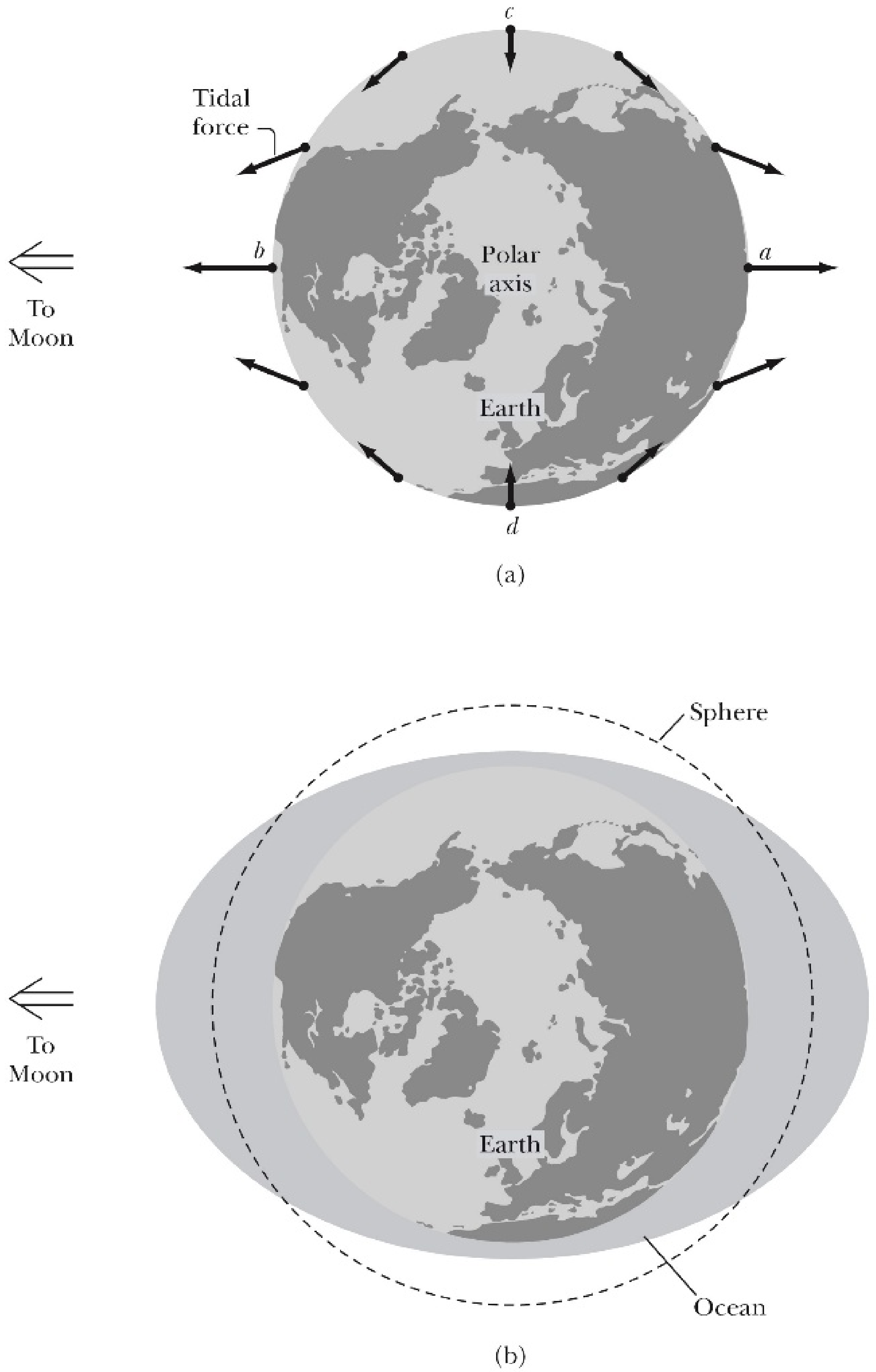
3. Point c: The x components of the vectors associated with R and D are similar, and the x components of the acceleration will cancel. The vector associated with D will have no y component, and the acceleration will be due to the y component associated with R (which points towards the center of the Earth):

$$\ddot{\vec{r}}_T = -GM_m\left(\frac{\hat{R}}{R^2} - \frac{\hat{D}}{D^2}\right) = -GM_m\frac{\hat{R}_y}{R^2} = -GM_m\frac{1}{D^2}\frac{r}{D}\hat{y} = -GM_m\frac{r}{D^3}\hat{y} \tag{5.28}$$

4. Point d: The calculation for the force at point d is similar to the calculation of the force at point c. The force will be directed towards the center of the Earth and will be equal to

$$\ddot{\vec{r}}_T = GM_m\frac{r}{D^3}\hat{y} \tag{5.29}$$

We thus conclude that the acceleration of the water on the surface of the Earth is directed away from the surface at two different locations. It will thus be high tide at these two locations and since the moon rotates around the Earth in one day, each location will see a high tide twice a day (one time when the moon that location is facing the moon, and one time when the moon is located on the opposite side of the Earth from that location). A summary of the tidal forces on the surface of the Earth is shown in Fig. 5.7.



(a)

(b)

Figure 5.7: Tidal forces at various places on the surface of the Earth.

6 Some Methods in the Calculus of Variations

In this Chapter, we focus on an important method of solving certain problems in Classical Mechanics. In many problems we need to determine how a system evolves between an initial state and a final state. For example, consider two locations on a two-dimensional plane. Consider an object moving from one location to the other, and assume the object experiences a friction force when it moves over the surface. One can ask questions such as "What is the path between the initial and the final condition that minimizes the work done by the friction force?". Such questions can be most easily answered using the calculus of variations. The evolution that can be studied using the calculus of variations is not limited to evolutions in real space. For example, one can consider the evolution of a gas in a pV diagram and ask "What is the path between the initial and final state that maximizes the work done by the gas?"

6.1 Euler's Equation

Consider the two-dimensional plane shown in Fig. 6.1. Our initial position is specified by (x_1,y_1) and the final position is specified by (x_2,y_2). We are asked to minimize the path integral of a function f between position 1 and position 2. Suppose the path integral is minimized when we use path $y(x)$. If we change the path slightly by adding a second function $n(x)$ then we expect that the path integral to increase. Consider the following path:

$$y(\alpha,x) = y(0,x) + \alpha\eta(x) \tag{6.1}$$

The function $\eta(x)$ is an arbitrary function of x and is used to make small changes to the path. The only requirements of $\eta(x)$ are that $\eta(x)$ has a continuous first derivative and that $\eta(x)$ vanishes at the end points of the path, that is $\eta(x_1) = \eta(x_2) = 0$. Since the path integral is minimized when we follow the path, the path integral must have an extreme value when $\alpha = 0$. This requires that

$$\left.\frac{\partial J}{\partial \alpha}\right|_{\alpha=0} = \left.\frac{\partial}{\partial \alpha}\int_{x_1}^{x_2} f\left(y(\alpha,x), y'(\alpha,x); x\right) dx\right|_{\alpha=0} = 0 \tag{6.2}$$

The left-hand side of this equation can be rewritten by differentiating the argument of the integral with respect to α:

$$\left.\int_{x_1}^{x_2}\left(\frac{\partial f}{\partial y}\frac{\partial y}{\partial \alpha} + \frac{\partial f}{\partial y'}\frac{\partial y'}{\partial \alpha}\right) dx\right|_{\alpha=0} = 0 \tag{6.3}$$

Using our definition of $y(\alpha, x)$ we can easily show that

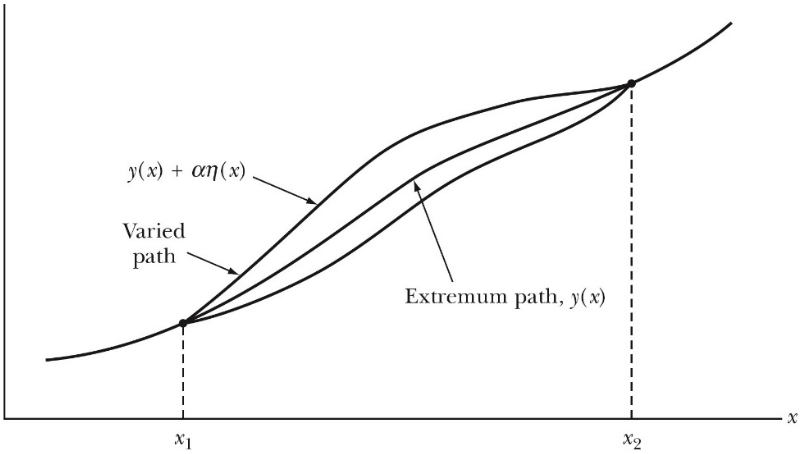


Figure 6.1: Path $y(x)$ that is used to move from position 1 to position 2.

$$\frac{\partial y}{\partial \alpha} = \frac{\partial}{\partial \alpha}\left(y\left(0,x\right)+\alpha\eta\left(x\right)\right) = \eta\left(x\right) \tag{6.4}$$

and

$$\frac{\partial y'}{\partial \alpha} = \frac{\partial}{\partial \alpha}\frac{dy}{dx} = \frac{\partial}{\partial \alpha}\left(\frac{dy\left(0,x\right)}{dx}+\alpha\frac{d\eta\left(x\right)}{dx}\right) = \frac{d\eta\left(x\right)}{dx} \tag{6.5}$$

The requirement that the path integral has an extreme is equivalent to requiring that

$$\int_{x_1}^{x_2}\left(\frac{\partial f}{\partial y}\eta\left(x\right)+\frac{\partial f}{\partial y'}\frac{d\eta\left(x\right)}{dx}\right)dx\Big|_{\alpha=0} = 0 \tag{6.6}$$

In order to simplify this equation we use the following relation to rewrite the second time in the integrant:

$$\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\eta(x)\right) = \left(\frac{\partial f}{\partial y'}\frac{d\eta}{dx}\right)+\left[\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right]\eta(x) \tag{6.7}$$

or

$$\left(\frac{\partial f}{\partial y'}\frac{d\eta}{dx}\right) = \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\eta(x)\right)-\left[\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right]\eta(x) \tag{6.8}$$

The second term in Eq. 6.6 can now be rewritten as

$$\begin{aligned}\int_{x_1}^{x_2}\left(\frac{\partial f}{\partial y'}\frac{d\eta(x)}{dx}\right)dx\Big|_{\alpha=0} &= \int_{x_1}^{x_2}\left(\frac{\partial f}{\partial y'}d\eta\left(x\right)\right)\Big|_{\alpha=0} = \\ &= \frac{\partial f}{\partial y'}\eta\left(x\right)\Big|_{x_1}^{x_2}-\int_{x_1}^{x_2}\left(\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\eta\left(x\right)\right)dx = \\ &= -\int_{x_1}^{x_2}\left(\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\eta\left(x\right)\right)dx\end{aligned} \tag{6.9}$$

The path integral is an extreme if

$$\int_{x_1}^{x_2}\left(\frac{\partial f}{\partial y}\eta\left(x\right)-\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\eta\left(x\right)\right)dx\Big|_{\alpha=0} = \int_{x_1}^{x_2}\left(\frac{\partial f}{\partial y}-\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right)\eta\left(x\right)dx\Big|_{\alpha=0} = 0 \tag{6.10}$$

Since $\eta(x)$ is an arbitrary function, this equation can only satisfied if

$$\frac{\partial f}{\partial y}-\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right) = 0 \tag{6.11}$$

Equation 6.11 is known as **Euler's equation**.

6.1.1 Example: Problem 6.4

Show that the geodesic on the surface of a right-circular cylinder is a segment of a helix.

The element of distance along the surface of a cylinder is

$$dS = \sqrt{\left(dx\right)^2+\left(dy\right)^2+\left(dz\right)^2} \tag{6.12}$$

In cylindrical coordinates, (x, y, z) is related to (ρ, ϕ, z) by

$$\left.\begin{aligned} x &= \rho\cos\phi \\ y &= \rho\sin\phi \\ z &= z \end{aligned}\right] \tag{6.13}$$

where ρ is the radius of the cylinder, which is constant.
Since we consider motion on the surface of a cylinder, the radius r is constant. The expression for x, y, and z can be used to express dx, dy, and dz in cylindrical coordinates:

$$\left.\begin{aligned} dx &= -\rho\sin\phi d\phi \\ dy &= \rho\cos\phi d\phi \\ dz &= dz \end{aligned}\right] \tag{6.14}$$

Substituting Eq. Ref. (6.14) into Ref. (6.12) and integrating along the entire path, we find

$$S = \int_1^2 \sqrt{\rho^2 \ (d\phi)^2 + (dz)^2} = \int_{\phi_1}^{\phi_2} \sqrt{\rho^2 + \left(\frac{dz}{d\phi}\right)^2} \ d\phi = \int_{\phi_1}^{\phi_2} \sqrt{\rho^2 + \dot{z}^2} \ d\phi \tag{6.15}$$

If S is to be a minimum, $f \equiv \sqrt{\rho^2 + \dot{z}^2}$ must satisfy the Euler equation:

$$\frac{\partial f}{\partial z} - \frac{\partial}{\partial \phi}\frac{\partial f}{\partial \dot{z}} = 0 \tag{6.16}$$

Since $\frac{\partial f}{\partial z} = 0$, the Euler equation becomes

$$\frac{\partial}{\partial \phi}\frac{\dot{z}}{\sqrt{\rho^2 + \dot{z}^2}} = 0 \tag{6.17}$$

This condition will be satisfied if

$$\frac{\dot{z}}{\sqrt{\rho^2 + \dot{z}^2}} = \text{constant} \equiv C \tag{6.18}$$

or

$$\dot{z} = \sqrt{\frac{C^2}{1 - C^2}}\rho \tag{6.19}$$

Since ρ is constant, Ref. (6.19) implies that

$$\frac{dz}{d\phi} = \text{constant} \tag{6.20}$$

and for any point along the path, z and ϕ change at the same rate. The curve described by this condition is a *helix*.

6.1.2 Example: Problem 6.7

Consider light passing from one medium with index of refraction n_1 into another medium with index of refraction n_2 (see Fig. 6.2). Using Fermat's principle to minimize time, derive the law of refraction: $n_1 \sin \theta_1 = n_2 \sin \theta_2$.

The time to travel the path shown in Fig. 6.2 is

$$t = \int \frac{ds}{v} = \int \frac{\sqrt{(dx)^2 + (dy)^2}}{v} = \int \frac{\sqrt{1 + \left(\frac{dy}{dx}\right)^2}}{v} dx = \int \frac{\sqrt{1 + y'^2}}{v} \ dx \tag{6.21}$$

The velocity v is v_1 when $x > 0$ and v_2 when $x < 0$. The function f is given by

$$f(y, y'; x) = \frac{\sqrt{1 + y'^2}}{v} \tag{6.22}$$

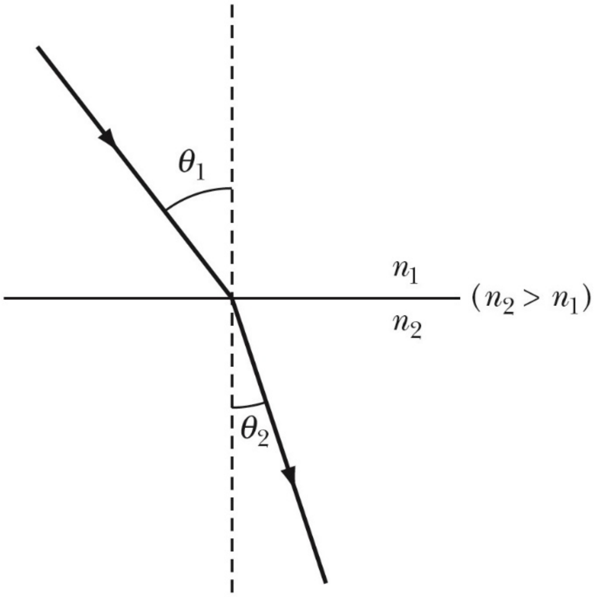


Figure 6.2: Problem 6.7. We will use the dashed line as our x axis and the interface between the two media as our y axis.

Since the velocity is not a function of y, $\partial f/\partial y = 0$. The Euler equation tells us

$$\frac{\partial f}{\partial y} - \frac{d}{dx}\left[\frac{\partial f}{\partial y'}\right] = -\frac{d}{dx}\left[\frac{y'}{v\sqrt{1+y'^2}}\right] = 0 \tag{6.23}$$

Equation 6.23 requires that

$$\left[\frac{y'}{v\sqrt{1+y'^2}}\right] = \text{constant} \tag{6.24}$$

Looking at Fig. 6.2 we find that $y = x\tan\theta$ and $y' = \tan\theta$. Using $v = c/n$ we can rewrite Eq. 6.24 as

$$\frac{y'}{v\sqrt{1+y'^2}} = \frac{\tan\theta}{\left(\frac{c}{n}\right)\sqrt{1+\tan^2\theta}} = \frac{\tan\theta}{\left(\frac{c}{n}\right)\sqrt{1+\tan^2\theta}} = \frac{\tan\theta}{\left(\frac{c}{n}\right)\sqrt{1+\frac{\sin^2\theta}{\cos^2\theta}}} = \left(\frac{n}{c}\right)\sin\theta = \text{constant} \tag{6.25}$$

This proves the assertion.

6.2 Second Form of Euler's Equation

In some applications, the function f may not depend explicitly on x: $\partial f/\partial x = 0$. In order to benefit from this constraint, it would be good to try to rewrite Euler's equation with a term $\partial f/\partial x$ instead of a term $\partial f/\partial y$.
The first step in this process is to examine df/dx. The general expression for df/dx is

$$\frac{df}{dx} = \frac{d}{dx}f\left(y, y' : x\right) = \frac{\partial f}{\partial y}\frac{dy}{dx} + \frac{\partial f}{\partial y'}\frac{dy'}{dx} + \frac{\partial f}{\partial x} = y'\frac{\partial f}{\partial y} + y''\frac{\partial f}{\partial y'} + \frac{\partial f}{\partial x} \tag{6.26}$$

Note that we have not assumed that $\partial f/\partial x = 0$ in order to derive this expression for df/dx. This equation can be rewritten as

$$y'' \frac{\partial f}{\partial y'} = \frac{df}{dx} - \frac{\partial f}{\partial x} - y' \frac{\partial f}{\partial y} \tag{6.27}$$

We also know that

$$\frac{d}{dx}\left(y' \frac{\partial f}{\partial y'} \right) = y'' \frac{\partial f}{\partial y'} + y' \frac{d}{dx} \frac{\partial f}{\partial y'} = \left(\frac{df}{dx} - y' \frac{\partial f}{\partial y} - \frac{\partial f}{\partial x} \right) + y' \frac{d}{dx} \frac{\partial f}{\partial y'} = \frac{df}{dx} - \frac{\partial f}{\partial x} + y' \left(\frac{d}{dx} \frac{\partial f}{\partial y'} - \frac{\partial f}{\partial y} \right) \tag{6.28}$$

Applying Euler's theorem to the term in the parenthesis on the right hand side, we can rewrite this equation as

$$\frac{d}{dx}\left(y' \frac{\partial f}{\partial y'} \right) = \frac{df}{dx} - \frac{\partial f}{\partial x} \tag{6.29}$$

or

$$\frac{\partial f}{\partial x} - \frac{d}{dx}\left(f - y' \frac{\partial f}{\partial y'} \right) = 0 \tag{6.30}$$

This equation is called **the second form of Euler's equation**. If f does not depend explicitly on x, $\partial f / \partial x = 0$, we conclude that

$$f - y' \frac{\partial f}{\partial y'} = \text{constant} \tag{6.31}$$

6.2.1 Example: Problem 6.4 – Part II

Show that the geodesic on the surface of a right-circular cylinder is a segment of a helix.

We have already solved this problem using the "normal" form of Euler's equation. However, looking back at the solution, we realize that the expression for f, $f \equiv \sqrt{\rho^2 + \dot{z}^2}$, does not depend explicitly on ϕ. We thus should also be able to use the second form of Euler's equation to solve this problem.

$$f - \dot{z} \frac{\partial f}{\partial \dot{z}} = \sqrt{\rho^2 + \dot{z}^2} - \dot{z} \frac{\dot{z}}{\sqrt{\rho^2 + \dot{z}^2}} = \frac{\rho^2}{\sqrt{\rho^2 + \dot{z}^2}} = \text{constant} = C \tag{6.32}$$

This equation requires that

$$\dot{z} = \frac{dz}{d\phi} = \sqrt{\frac{\rho^4}{C^2} - \rho^2} = \rho \sqrt{\frac{\rho^2}{C^2} - 1} = \text{constant} \tag{6.33}$$

Since ρ is constant, this equation implies that

$$\frac{dz}{d\phi} = \text{constant} \tag{6.34}$$

and for any point along the path, z and ϕ change at the same rate. The curve described by this condition is a *helix*.

6.3 Euler's Equation with Several Dependent Variables

Consider a situation where the function f depends on several dependent variables $y_1, y_2, y_3, \ldots$ etc., each of which depends on the independent variable x. Each dependent variable $y_i(\alpha, x)$ is related to the solution $y_i(0, x)$ in the following manner:

$$y_i(\alpha, x) = y_i(0, x) + \alpha \eta_i(x) \tag{6.35}$$

If the dependent variables y_1, y_2, y_3, etc. minimize the path integral of f, they must satisfy the following condition:

$$\frac{\partial f}{\partial y_i} - \frac{d}{dx}\left(\frac{\partial f}{\partial {y_i}'}\right) = 0 \tag{6.36}$$

for $i = 1, 2, 3, \ldots$. The procedure to find the optimum paths is similar to the procedures we have discussed already, except that we need to solve the Euler equation for each dependent variable y_i.

6.4 Euler's Equation with Boundary Conditions

In many cases, the dependent variable y must satisfy certain boundary conditions. For example, in problem 6.4 the function y must be located on the surface of the cylinder. In this case, any point on y must satisfy the following condition:

$$r = \rho = \text{constant} \tag{6.37}$$

In general, we can specify the constraint on the path(s) by using one or more functions g and requiring that $g\{y_i; x\} = 0$.
Let us start with the case where we have two dependent variables y and z. In this case, we can write the function f as

$$f = f\{y, y', z, z'; x\} \tag{6.38}$$

In this case, we can write the differential of J with respect to α as

$$\frac{\partial J}{\partial \alpha} = \int_{x_1}^{x_2} \left\{\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right)\eta_y(x) + \left(\frac{\partial f}{\partial z} - \frac{d}{dx}\left(\frac{\partial f}{\partial z'}\right)\right)\eta_z(x)\right\} dx \tag{6.39}$$

For our choice of constraint, we can immediately see that the derivative of g must be zero:

$$\frac{dg}{d\alpha} = \left(\frac{\partial g}{\partial y}\right)\frac{\partial y}{\partial \alpha} + \left(\frac{\partial g}{\partial z}\right)\frac{\partial z}{\partial \alpha} = 0 \tag{6.40}$$

Using our definition of y and z we can rewrite this equation as

$$\frac{dg}{d\alpha} = \left(\frac{\partial g}{\partial y}\right)\eta_y + \left(\frac{\partial g}{\partial z}\right)\eta_z = 0 \tag{6.41}$$

This equation shows us a general relation between the functions η_z and η_z:

$$\frac{\eta_z}{\eta_y} = -\frac{\left(\frac{\partial g}{\partial y}\right)}{\left(\frac{\partial g}{\partial z}\right)} \tag{6.42}$$

Using this relation, we can rewrite our expression for the differential of J:

$$\begin{aligned}
\frac{\partial J}{\partial \alpha} &= \int_{x_1}^{x_2} \left\{\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right)\eta_y(x) + \left(\frac{\partial f}{\partial z} - \frac{d}{dx}\left(\frac{\partial f}{\partial z'}\right)\right)\eta_z(x)\right\} dx = \\
&= \int_{x_1}^{x_2} \left\{\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right) + \left(\frac{\partial f}{\partial z} - \frac{d}{dx}\left(\frac{\partial f}{\partial z'}\right)\right)\frac{\eta_z(x)}{\eta_y(x)}\right\}\left\{\eta_y(x)\right\} dx \\
&= \int_{x_1}^{x_2} \left\{\left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right) - \left(\frac{\partial f}{\partial z} - \frac{d}{dx}\left(\frac{\partial f}{\partial z'}\right)\right)\frac{\left(\frac{\partial g}{\partial y}\right)}{\left(\frac{\partial g}{\partial z}\right)}\right\}\left\{\eta_y(x)\right\} dx
\end{aligned} \tag{6.43}$$

Since the function η_y is an arbitrary function, the equation can only evaluate to 0 if the term in the brackets is equal to 0:

$$\left(\frac{\partial f}{\partial y}-\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right)-\left(\frac{\partial f}{\partial z}-\frac{d}{dx}\left(\frac{\partial f}{\partial z'}\right)\right)\frac{\left(\frac{\partial g}{\partial y}\right)}{\left(\frac{\partial g}{\partial z}\right)}=0 \tag{6.44}$$

This equation can be rewritten as

$$\left(\frac{\partial f}{\partial y}-\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right)\left(\frac{\partial g}{\partial y}\right)^{-1}=\left(\frac{\partial f}{\partial z}-\frac{d}{dx}\left(\frac{\partial f}{\partial z'}\right)\right)\left(\frac{\partial g}{\partial z}\right)^{-1} \tag{6.45}$$

This equation can only be correct if both sides are equal to a function that depends only on x:

$$\left(\frac{\partial f}{\partial y}-\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right)\left(\frac{\partial g}{\partial y}\right)^{-1}=\left(\frac{\partial f}{\partial z}-\frac{d}{dx}\left(\frac{\partial f}{\partial z'}\right)\right)\left(\frac{\partial g}{\partial z}\right)^{-1}=-\lambda(x) \tag{6.46}$$

or

$$\left(\frac{\partial f}{\partial y}-\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right)+\lambda(x)\left(\frac{\partial g}{\partial y}\right)=0 \tag{6.47}$$

$$\left(\frac{\partial f}{\partial z}-\frac{d}{dx}\left(\frac{\partial f}{\partial z'}\right)\right)+\lambda(x)\left(\frac{\partial g}{\partial z}\right)=0 \tag{6.48}$$

Note that in this case, where we have one auxiliary condition, $g\{y_, z; x\} = 0$, we end up with one Lagrange undetermined multiplier $\lambda(x)$. Since we have three equations and three unknown (y, z, and λ), we can determine the unknown.
In certain problems, the constraint can only be written in integral form. For example, the constraint for problems dealing with ropes will be that the total length of the path is equal to the length of the rope L:

$$K[y]=\int g\{y,y';x\}\,dx=L \tag{6.49}$$

The curve y then must satisfy the following differential equation:

$$\left(\frac{\partial f}{\partial y}-\frac{d}{dx}\left(\frac{\partial f}{\partial y'}\right)\right)+\lambda(x)\left(\frac{\partial g}{\partial y}-\frac{d}{dx}\left(\frac{\partial g}{\partial y'}\right)\right)=0 \tag{6.50}$$

6.4.1 Example: Problem 6.12

Repeat example 6.4, finding the shortest path between any two points on the surface of a sphere, but use the method of the Euler equation with an auxiliary condition imposed.

The path length is given by

$$s=\int ds=\int\sqrt{1+y'^2+z'^2}\;dx \tag{6.51}$$

The function f is thus equal to

$$f(y,y',z,z';x)=\sqrt{1+y'^2+z'^2} \tag{6.52}$$

and our equation of constraint is

$$g(x,y,z)=x^2+y^2+z^2-\rho^2=0 \tag{6.53}$$

The Euler equations with undetermined multipliers, Eq. 6.47 and Eq. 6.48, tell us that

$$\frac{d}{dx}\left[\frac{y'}{\sqrt{1+y'^2+z'^2}}\right]=\lambda\frac{dg}{dy}=2\lambda y \tag{6.54}$$

and

$$\frac{d}{dx}\left[\frac{z'}{\sqrt{1+y'^2+z'^2}}\right]=\lambda\frac{dg}{dz}=2\lambda z \tag{6.55}$$

Eliminating the factor λ, we obtain

$$\frac{1}{y}\frac{d}{dx}\left[\frac{y'}{\sqrt{1+y'^2+z'^2}}\right]-\frac{1}{z}\frac{d}{dx}\left[\frac{z'}{\sqrt{1+y'^2+z'^2}}\right]=0 \tag{6.56}$$

This simplifies first to

$$z\left[y''\left(1+y'^2+z'^2\right)-y'(y'y''+z'z'')\right]-y\left[z''\left(1+y'^2+z'^2\right)-z'(y'y''+z'z'')\right]=0 \tag{6.57}$$

and then to

$$zy''+(yy'+zz')z'y''-yz''-(yy'+zz')y'z''=0 \tag{6.58}$$

Using the derivative of Ref. (6.53) we can rewrite Eq. 6.58 as

$$(z-xz')y''=(y-xy')z'' \tag{6.59}$$

This looks to be in the simplest form we can make it, but is it a plane? Take the equation of a plane passing through the origin:

$$Ax+By=z \tag{6.60}$$

and make it a differential equation by taking derivatives (giving $A+By'=z'$ and $by''=z''$ and eliminating the constants:

$$B=\frac{z''}{y''} \tag{6.61}$$

$$A=z'-By'=z'-\frac{z''}{y''} \tag{6.62}$$

This substitution yields Ref. (6.59) exactly. This confirms that the path must be the intersection of the sphere with a plane passing through the origin, as required.

6.4.2 Example: Problem 6.4 – Part III

Show that the geodesic on the surface of a right circular cylinder is a segment of a helix.

We have already solved this problem using the “normal” form of Euler’s equation. However, this problem is a good example of how to approach problems with constraints. **Note: doing it in this way is NOT easier than the approaches we have used previously.**
Let us consider two points on the surface of the cylinder:

$$x_i=(\rho\cos\phi_i,\rho\sin\phi_i,z_i) \tag{6.63}$$

and

$$x_f = \left(\rho\cos\phi_f, \rho\sin\phi_f, z_f\right) \tag{6.64}$$

Consider an arbitrary path connecting the initial and final position. The length of a tiny segment of this path is

$$dl = \sqrt{(dx)^2 + (dy)^2 + (dz)^2} = dz\sqrt{1 + \left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1} \tag{6.65}$$

The integral we want to minimize is

$$J = \int dl = \int \left(\sqrt{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1} \right) dz = \int \left(\sqrt{(x')^2 + (y')^2 + 1} \right) dz \tag{6.66}$$

We immediately see that z is our independent variable and that x, and y are our dependent variables. The function f is thus given by

$$f\,(x, x', y, y'; z) = \sqrt{(x')^2 + (y')^2 + 1} \tag{6.67}$$

The solution y is constrained to be on the surface of the cylinder, and the following equation of constraint needs to be applied:

$$g\,(x, y) = x^2 + y^2 - \rho^2 = 0 \tag{6.68}$$

For this equation of constraint we know that

$$\frac{\partial g}{\partial x} = 2x \tag{6.69}$$

and

$$\frac{\partial g}{\partial y} = 2y \tag{6.70}$$

Note: if we had picked our function of constraint to be

$$g\,(x, y) = \sqrt{x^2 + y^2} - \rho = 0 \tag{6.71}$$

we would get more complicated partial derivatives of g.

To solve the current problem, we thus need to solve the following Euler equations:

$$\left(\frac{\partial f}{\partial x} - \frac{d}{dz}\left(\frac{\partial f}{\partial x'}\right)\right) + \lambda\,(z)\left(\frac{\partial g}{\partial x}\right) = -\frac{d}{dz}\left(\frac{x'}{\sqrt{(x')^2 + (y')^2 + 1}}\right) + 2\lambda\,(z)\,x = 0 \tag{6.72}$$

$$\left(\frac{\partial f}{\partial y} - \frac{d}{dz}\left(\frac{\partial f}{\partial y'}\right)\right) + \lambda\,(z)\left(\frac{\partial g}{\partial y}\right) = -\frac{d}{dz}\left(\frac{y'}{\sqrt{(x')^2 + (y')^2 + 1}}\right) + 2\lambda\,(z)\,y = 0 \tag{6.73}$$

These two equations can be rewritten as

$$y\frac{d}{dz}\left(\frac{x'}{\sqrt{(x')^2 + (y')^2 + 1}}\right) = 2\lambda\,(z)\,xy \tag{6.74}$$

$$x\frac{d}{dz}\left(\frac{y'}{\sqrt{(x')^2 + (y')^2 + 1}}\right) = 2\lambda\,(z)\,xy \tag{6.75}$$

Eliminating λ we find that

$$y\frac{d}{dz}\left(\frac{x'}{\sqrt{(x')^2+(y')^2+1}}\right)=x\frac{d}{dz}\left(\frac{y'}{\sqrt{(x')^2+(y')^2+1}}\right) \tag{6.76}$$

This equation can be rewritten as

$$y\left(\frac{x''}{\sqrt{(x')^2+(y')^2+1}}-\frac{x'\{x'x''+y'y''\}}{\left((x')^2+(y')^2+1\right)^{3/2}}\right)=x\left(\frac{y''}{\sqrt{(x')^2+(y')^2+1}}-\frac{y'\{x'x''+y'y''\}}{\left((x')^2+(y')^2+1\right)^{3/2}}\right) \tag{6.77}$$

After simplifying this equation we obtain

$$y\left(x''\left(1+(y')^2\right)-x'y'y''\right)=x\left(y''\left(1+(x')^2\right)-x'y'x''\right) \tag{6.78}$$

Is this equation describing a helix? Yes, it is! How do you see that? Let us look at the definition of a helix:

$$\begin{aligned} x &= \rho\cos\phi \\ y &= \rho\sin\phi \\ z &= \beta\phi \end{aligned} \tag{6.79}$$

We find that

$$\begin{aligned} x' &= \tfrac{dx}{dz} = \tfrac{dx}{d\phi}\tfrac{d\phi}{dz} = \tfrac{1}{\beta}(-\rho\sin\phi) = -\tfrac{1}{\beta}y \\ y' &= \tfrac{dy}{dz} = \tfrac{dy}{d\phi}\tfrac{d\phi}{dz} = \tfrac{1}{\beta}(\rho\cos\phi) = \tfrac{1}{\beta}x \end{aligned} \tag{6.80}$$

and

$$\begin{aligned} x'' &= -\tfrac{1}{\beta}y' = -\tfrac{1}{\beta^2}x \\ y'' &= \tfrac{1}{\beta}x' = -\tfrac{1}{\beta^2}y \end{aligned} \tag{6.81}$$

Taking these relations and substituting them in the solution shown in Eq. 6.78 we find:

$$y\left(x''\left(1+(y')^2\right)-x'y'y''\right)=-\frac{1}{\beta^2}xy-\frac{1}{\beta^4}x^3y-\frac{1}{\beta^4}xy^3=x\left(y''\left(1+(x')^2\right)-x'y'x''\right) \tag{6.82}$$

6.5 The δ notation

It is common to use the δ notation in the calculus of variations. In order to use the δ notation, we use the following definitions:

$$\delta y=\frac{\partial y}{\partial\alpha}d\alpha \tag{6.83}$$

$$\delta J=\frac{\partial J}{\partial\alpha}d\alpha \tag{6.84}$$

In terms of these variables, we find

$$\begin{aligned}
\delta J &= \frac{\partial J}{\partial \alpha} d\alpha = \left\{ \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'} \right) \right) \frac{\partial y}{\partial \alpha} dx \right\} d\alpha = \\
&= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'} \right) \right) \frac{\partial y}{\partial \alpha} d\alpha dx = \\
&= \int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx}\left(\frac{\partial f}{\partial y'} \right) \right) \delta y dx
\end{aligned} \tag{6.85}$$

Since δy is an arbitrary function, the requirement that $\delta J = 0$ requires that the term in parentheses be 0. This, of course, is the Euler equation we have encountered before!

It is important to note that there is a significant difference between δy and dy. Based on the definition of δy we see that δy tells us how y varies when we change α while keeping all other variables fixed (including, for example, time t).

7 Hamilton's Principle - Lagrangian and Hamiltonian Dynamics

Many interesting physics systems describe systems of particles on which many forces are acting. Some of these forces are immediately obvious to the person studying the system since they are externally applied. Other forces are not immediately obvious, and are applied by the external constraints imposed on the system. These forces are often difficult to quantify, but the effect of these forces is easy to describe. Trying to describe such a system in terms of Newton's equations of motion is often difficult since it requires us to specify the total force. In this Chapter we will see that describing such a system by applying Hamilton's principle will allow us to determine the equation of motion for systems for which we would not be able to derive these equations easily on the basis of Newton's laws. We should stress however, that Hamilton's principle does not provide us with a new physical theory, but it allows us to describe the existing theories in a new and elegant framework.

7.1 Hamilton's Principle

The evolution of many physical systems involves the minimization of certain physical quantities. We already have encountered an example of such a system, namely the case of refraction where light will propagate in such a way that the total time of flight is minimized. This same principle can be used to explain the law of reflection: the angle of incidence is equal to the angle of reflection.

The minimization approach to physics was formalized in detail by Hamilton, and resulted in Hamilton's Principle which states:

" Of all the possible paths along which a dynamical system may move from one point to another within a specified time interval (consistent with any constraints), the actual path followed is that which minimizes the time integral of the difference between the kinetic and potential energies. "

We can express this principle in terms of the calculus of variations:

$$\delta \int_{t_1}^{t_2} (T - U)\, dt = 0 \tag{7.1}$$

The quantity T - U is called the Lagrangian L.

Consider first a single particle, moving in a conservative force field. For such a particle, the kinetic energy T will just be a function of the velocity of the particle, and the potential energy will just be a function of the position of the particle. The Lagrangian is thus also a function of the position and the velocity of the particle. Hamilton's theorem states that we need to minimize the Lagrangian and thus require that

$$\delta \int_{t_1}^{t_2} L(x_i, \dot{x}_i)\, dt = 0 \tag{7.2}$$

In Chapter 6 we have developed the theory required to solve problems of this type and found that the Lagrangian must satisfy the following relation:

$$\frac{\partial L}{\partial x_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}_i} = 0 \tag{7.3}$$

This last equation is called the **Lagrange equation of motion.** Note that in order to generate this equation of motion, we do not need to know the forces. Information about the forces is included in the details of the kinetic and potential energy of the system.

Consider the example of a plane pendulum. For this system, there is only one coordinate we need to specify, namely the polar angle θ. The kinetic energy T of the pendulum is equal to

$$T = \frac{1}{2} m l^2 \dot{\theta}^2 \tag{7.4}$$

and the potential energy U is given by

$$U = mgl(1-\cos\theta) \tag{7.5}$$

The Lagrangian for this system is thus equal to

$$L = T - U = \frac{1}{2}ml^2\dot{\theta}^2 - mgl(1-\cos\theta) \tag{7.6}$$

The equation of motion can now be determined and is found to be equal to

$$\frac{\partial L}{\partial \theta} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} = -mgl\sin\theta - \frac{d}{dt}\left(ml^2\dot{\theta}\right) = -mgl\sin\theta - ml^2\ddot{\theta} = 0 \tag{7.7}$$

or

$$\ddot{\theta} + \frac{g}{l}\sin\theta = 0 \tag{7.8}$$

This equation is of course the same equation we can find by applying Newton's force laws. In this example, the only coordinate that was used was the polar angle θ. Even though the pendulum is a 3-dimensional system, the constraints imposed upon its motion reduced the number of degrees of freedom from 3 to 1.

7.2 Generalized Coordinates

If we try to describe a system of n particles, we generally need $3n$ coordinates to specify the position of its components. If external constraints are imposed on the system, the number of degrees of freedom may be less. If there are m constraints applied, the number of degrees of freedom will be $3n$ - m. The coordinates do not need to be the coordinates of a coordinate system, but can be any set of quantities that completely specify the state of the system. The state of the system is thus fully specified by a point in the configuration space (which is a $3n$ - m dimensional space). The time evolution of the system can be described by a path in the configuration space.

The generalized coordinates of a system are written as q_1, q_2, q_3, ... The generalized coordinates are of course related to the physical coordinates of the particles being described:

$$x_{\alpha,i} = x_{\alpha,i}(q_1, q_2, q_3, ..., t) = x_{\alpha,i}\left(q_j, t\right) \tag{7.9}$$

where i = 1, 2, 3 and α = 1, 2, ..., n. Since the generalized coordinates in general will depend on time, we can also introduce the generalized velocities. The physical velocities will depend on the generalized velocities:

$$\dot{x}_{\alpha,i} = \dot{x}_{\alpha,i}\left(q_j, \dot{q}_j, t\right) \tag{7.10}$$

7.3 Equations of Motion in Generalized Coordinates

Based on the introduction of the Lagrangian and generalized coordinates, we can rephrase Hamilton's principle in the following way:

" Of all the possible paths along which a dynamical system may move from one point to another in configuration space within a specified time interval (consistent with any constraints), the actual path followed is that which minimizes the time integral of the Lagrangian function for the system. "

Thus

$$\delta\int_{t_1}^{t_2} L(q_i, \dot{q}_i, t)\,dt = 0 \tag{7.11}$$

and

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = 0 \tag{7.12}$$

When we use the Lagrange's equations to describe the evolution of a system, we must recognize that these equations are only correct when the following conditions are met:

1. The force(s) acting on the system, except the forces of constraint, must be derivable from one or more potentials.
2. The equations of constraint must be relations that connect the coordinates of the particles, and may be time-dependent (note: this means that they are independent of velocity).

Constraints that do not depend on the velocities are called **holonomic constraints**. There are two different types of holonomic constraints:

1. Fixed or **scleronomic constraints**: constraints that do not depend on time.
2. Moving or **rheonomic constraints**: constraints that depend on time.

7.3.1 Example: Problem 7.4

A particle moves in a plane under the influence of a force f = -Ar^{a-1} directed toward the origin; A and α are constants. The coordinates used to describe this motion are shown in Fig. 7.1. Choose appropriate generalized coordinates, and let the potential energy be zero at the origin. Find the Lagrangian equations of motion. Is the angular momentum about the origin conserved? Is the total energy conserved?

If we choose (r,θ) as the generalized coordinates, the kinetic energy of the particle is

$$T = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) \tag{7.13}$$

Since the force is related to the potential by

$$f = -\frac{\partial U}{\partial r} \tag{7.14}$$

we find

$$U = \frac{A}{\alpha} r^{\alpha} \tag{7.15}$$

where we let $U(r = 0) = 0$. Therefore, the Lagrangian becomes

$$L = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - \frac{A}{\alpha} r^{\alpha} \tag{7.16}$$

Lagrange's equation for the coordinate r leads to

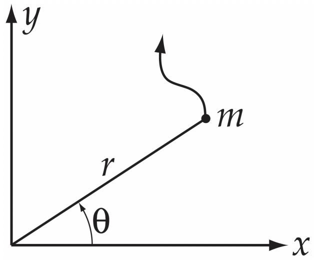


Figure 7.1: Problem 7.4

$$m\ddot{r} - mr\dot{\theta}^2 + Ar^{\alpha-1} = 0 \tag{7.17}$$

Lagrange's equation for the coordinate θ leads to

$$\frac{d}{dt}\left(mr^2\dot{\theta}\right) = 0 \tag{7.18}$$

Since $mr^2\dot{\theta} = \ell$ is identified as the angular momentum, (7.18) implies that angular momentum is conserved. Using ℓ, we can rewrite (7.17) as

$$m\ddot{r} - \frac{\ell^2}{mr^3} + Ar^{\alpha-1} = 0 \tag{7.19}$$

Multiplying Eq. (7.19) by $\dot{r}$, we have

$$m\dot{r}\ddot{r} - \frac{\dot{r}\ell^2}{mr^3} + A\dot{r}r^{\alpha-1} = 0 \tag{7.20}$$

which is equivalent to

$$\frac{d}{dt}\left[\frac{1}{2}m\dot{r}^2\right] + \frac{d}{dt}\left[\frac{\ell^2}{2mr^2}\right] + \frac{d}{dt}\left[\frac{A}{\alpha}r^{\alpha}\right] = 0 \tag{7.21}$$

Therefore,

$$\frac{d}{dt}(T + U) = 0 \tag{7.22}$$

and the total energy is conserved.

7.3.2 Example: Problem 7.8

Consider a region of space divided by a plane as shown in Fig. 7.2. The potential energy of a particle in region 1 is U_1 and in region 2 it is U_2. If a particle of mass m and speed v_1 in region 1 passes from region 1 to region 2 such that its path in region 1 makes an angle θ_1 with the normal to the plane of separation and an angle θ_2 with the normal when in region 2, show that

$$\frac{\sin\theta_1}{\sin\theta_2} = \sqrt{1 + \frac{U_1 - U_2}{T_1}} \tag{7.23}$$

where $T_1 = (1/2)m{v_1}^2$,

Let us choose the (x, y) coordinates so that the two regions are divided by the y axis:

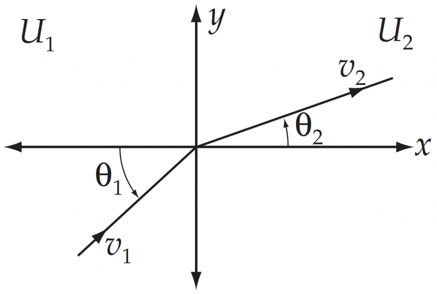


Figure 7.2: Problem 7.8

$$U(x) = \begin{bmatrix} U_1 & x < 0 \\ \\ U_2 & x > 0 \end{bmatrix} \tag{7.24}$$

Since the potential energy is a function of x, the Lagrangian of the particle can be written as

$$L = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2\right) - U(x) \tag{7.25}$$

Therefore, Lagrange's equations for the coordinates x and y are

$$m\ddot{x} + \frac{dU(x)}{dx} = 0 \tag{7.26}$$

$$m\ddot{y} = 0 \tag{7.27}$$

Using the relation

$$m\ddot{x} = \frac{d}{dt}m\dot{x} = \frac{dp_x}{dt} = \frac{dp_x}{dx}\frac{dx}{dt} = \frac{p_x}{m}\frac{dp_x}{dx} \tag{7.28}$$

Eq. (7.26) becomes

$$\frac{p_x}{m}\frac{dp_x}{dx} + \frac{dU(x)}{dx} = 0 \tag{7.29}$$

Integrating (7.29) from any point in region 1 to any point in region 2, we find

$$\int_1^2 \frac{p_x}{m}\frac{dp_x}{dx}dx + \int_1^2 \frac{dU(x)}{dx}dx = 0 \tag{7.30}$$

$$\frac{p_{x_2}^2}{2m} - \frac{p_{x_1}^2}{2m} + U_2 - U_1 = 0 \tag{7.31}$$

or, equivalently,

$$\frac{1}{2}m\dot{x}_1^2 + U_1 = \frac{1}{2}m\dot{x}_2^2 + U_2 \tag{7.32}$$

We can rewrite Eq. (7.27) as

$$\frac{d}{dt}m\dot{y} = 0 \tag{7.33}$$

and conclude that $m\dot{y}$ is constant. Therefore,

$$m\dot{y}_1 = m\dot{y}_2 \tag{7.34}$$

From (7.34) we have

$$\frac{1}{2}m\dot{y}_1^2 = \frac{1}{2}m\dot{y}_2^2 \tag{7.35}$$

Adding Eq. (7.32) and Eq. (7.35), we have

$$\frac{1}{2}mv_1^2 + U_1 = \frac{1}{2}mv_2^2 + U_2 \tag{7.36}$$

From Eq. (7.34) we also have

$$mv_1\sin\theta_1 = mv_2\sin\theta_2 \tag{7.37}$$

Substituting Eq. (7.36) into Eq. (7.37), we find

$$\frac{\sin\theta_1}{\sin\theta_2} = \frac{v_2}{v_1} = \left[1 + \frac{U_1 - U_2}{T_1}\right]^{1/2} \tag{7.38}$$

This problem is the mechanical analog of the refraction of light upon passing from a medium of a certain optical density into a medium with a different optical density.

7.4 Lagrange's Equations with Undetermined Multipliers

We have seen already a number of examples where one could remove the equations of constraint by a suitable choice of coordinates. For example, when we looked at the motion of an object on the surface of a cylinder we could either:

1. Use a set of three coordinates to describe the motion, coupled with one equation of constraint.
2. Use a set of two coordinates (such at the azimuthal angle and the vertical position) to describe the motion, without an equation of constraint.

In this Section we will look at situations where the constraint depends on the velocity:

$$f\left(x_{\alpha,i}, \dot{x}_{\alpha,i}, t\right) = 0 \tag{7.39}$$

If the constraints can be expressed in a differential form,

$$\sum_{j=1}^{s} \frac{\partial f_k}{\partial q_j} dq_j = 0 \tag{7.40}$$

we can directly incorporate it into the Lagrange equations:

$$\frac{\partial L}{\partial q_j} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_j} + \sum_{k=1}^{m} \lambda_k(t) \frac{\partial f_k}{\partial q_j} = 0 \tag{7.41}$$

It turns out that the forces of constraint can be determined from the equations of constraints and the Lagrange multipliers $\lambda_m(t)$:

$$Q_j = \sum_{k=1}^{m} \lambda_k(t) \frac{\partial f_k}{\partial q_j} \tag{7.42}$$

where Q_j is the jth component of the generalized force, expressed in terms of the generalized coordinates. The use of Lagrange multiplier to determine the forces of constraints is nicely illustrated in Example 7.9 in the textbook, where a disk rolling down an inclined plane is being studied. If the disk does not slip, we find that the distance along the plane y and the angle of rotation θ are related, and the equation of constraint is

$$f(y, \theta) = y - R\theta = 0 \tag{7.43}$$

The textbook explains in detail how the Lagrange equations are solved in this case, and I will not reproduce this here. The solution shows us that the Lagrange multiplier is given by

$$\lambda = -\frac{Mg\sin\alpha}{3} \tag{7.44}$$

By combining the equation of constraint and the Lagrange multipliers we can determine the generalized forces of constraint:

$$Q_y = \lambda(t)\frac{\partial f}{\partial y} = -\frac{Mg\sin\alpha}{3} \tag{7.45}$$

and

$$Q_\theta = \lambda(t)\frac{\partial f}{\partial \theta} = \frac{MgR\sin\alpha}{3} \tag{7.46}$$

Note that these forces of constraint do not have to be all pure forces. The force of constraint associated with y is the friction force between the disk and the plane that is required to ensure that the disk rolls without slipping. However, the force of constraint associated with the angle θ is the torque of this friction force with respect to the center of the disk. We need to note the generalized force does not have to have the unit of force.

It is also important to note that if we had chosen to solve the problem by expressing the Lagrangian in terms of a single coordinate y, by eliminating the angle, we would not have obtained any information about the forces of constraint. Although I have stressed that in many cases, you can simplify the solution by the proper choice of coordinates such that the equations of constraint are eliminated, in this case, the solution will not provide any information about the forces of constraint.

7.4.1 Example: Problem 7.12

A particle of mass m rests on a smooth plane. The plane is raised to an inclination angle θ at a constant rate α ($\theta = 0°$ at $t = 0$), causing the particle to move down the plane. Determine the motion of the particle.

This problem is an example of a problem with a velocity-dependent constraint. However, if we can easily incorporate the constraint into the Lagrangian, we do not need to worry about constraint functions. In this example, we use our knowledge of the constraint immediately in our expression of the kinetic and the potential energy. Putting the origin of our coordinate system at the bottom of the plane, as shown in Fig. 7.3, we find

$$L = T - U = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - mgr\sin\theta \tag{7.47}$$

$$\theta = \alpha t; \dot{\theta} = \alpha \tag{7.48}$$

$$L = \frac{1}{2}m\left(\dot{r}^2 + \alpha^2 r^2\right) - mgr\sin\alpha t \tag{7.49}$$

Lagrange's equation for r gives

$$m\ddot{r} = m\alpha^2 r - mg\sin\alpha t \tag{7.50}$$

or

$$\ddot{r} - \alpha^2 r = -g\sin\alpha t \tag{7.51}$$

The general solution is of the form $r = r_h + r_n$ where r_h is the general solution of the homogeneous equation $\ddot{r} - \alpha^2 r = 0$ and r_n is a particular solution of Eq. (7.51). So

$$r_h = Ae^{\alpha t} + Be^{-\alpha t} \tag{7.52}$$

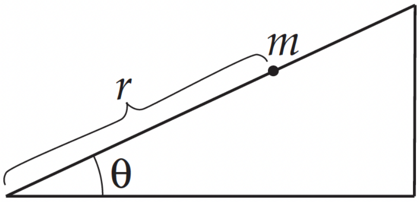


Figure 7.3: Problem 7.12

For r_n, try a solution of the form $r_p = C\,\sin at$. Then $\ddot{r}_p = -C\alpha^2\sin\alpha t$. Substituting into (7.51) gives

$$-C\alpha^2\sin\alpha t - C\alpha^2\sin\alpha t = -g\sin\alpha t \tag{7.53}$$

$$C = \frac{g}{2\alpha^2} \tag{7.54}$$

So

$$r(t) = Ae^{\alpha t} + Be^{-\alpha t} + \frac{g}{2\alpha^2}\sin\alpha t \tag{7.55}$$

We can determine A and B from the initial conditions:

$$r(0) = r_0 \tag{7.56}$$

$$\dot{r}(0) = 0 \tag{7.57}$$

Equation (7.56) implies:

$$r_0 = A + B \tag{7.58}$$

Equation (7.57) implies:

$$0 = A - B + \frac{g}{2\alpha^2} \tag{7.59}$$

Solving for A and B gives:

$$A = \frac{1}{2}\left[r_0 - \frac{g}{2\alpha^2}\right] B = \frac{1}{2}\left[r_0 + \frac{g}{2\alpha^2}\right] \tag{7.60}$$

$$r(t) = \frac{1}{2}\left[r_0 - \frac{g}{2\alpha^2}\right]e^{\alpha t} + \frac{1}{2}\left[r_0 + \frac{g}{2\alpha^2}\right]e^{-\alpha t} + \frac{g}{2\alpha^2}\sin\alpha t \tag{7.61}$$

or

$$r(t) = r_0\cosh\alpha t + \frac{g}{2\alpha^2}(\sin\alpha t - \sinh\alpha t) \tag{7.62}$$

Although we have found an analytical solution to this problem, we need to examine if the solution matches our expectation of the motion of the mass m. The best way to do this is to plot a graph of the motion of the mass in a Cartesian coordinate system. Consider the situation where $r_0 = 10$ m. Figure 7.4 shows the trajectory of the mass for two different value of α: $\alpha = 0.1$ rad/s and $\alpha = 0.03$ rad/s.

7.5 The generalized momentum

One of the big differences between the equations of motion obtained from the Lagrange equations and those obtained from Newton's equations is that in the latter case, the coordinate frame used is always a Cartesian coordinate frame. When we use the Lagrange equations we have the option to choose generalized coordinates that do not have to correspond to the coordinates of a Cartesian coordinate system.

The generalized coordinates are related to the Cartesian coordinates, and transformation rules allow use to carry out transformations between coordinate systems. The generalized forces of constraint are related to the Newtonian forces of constraint, as was illustrated in Example 7.9 in the textbook. The similarities between the Cartesian and the generalized parameters suggest it may also be useful to consider the concept of the generalized momentum.

In a Cartesian coordinate system we can easily determine the connection between the Lagrangian and the linear momentum. The Lagrangian is equal to

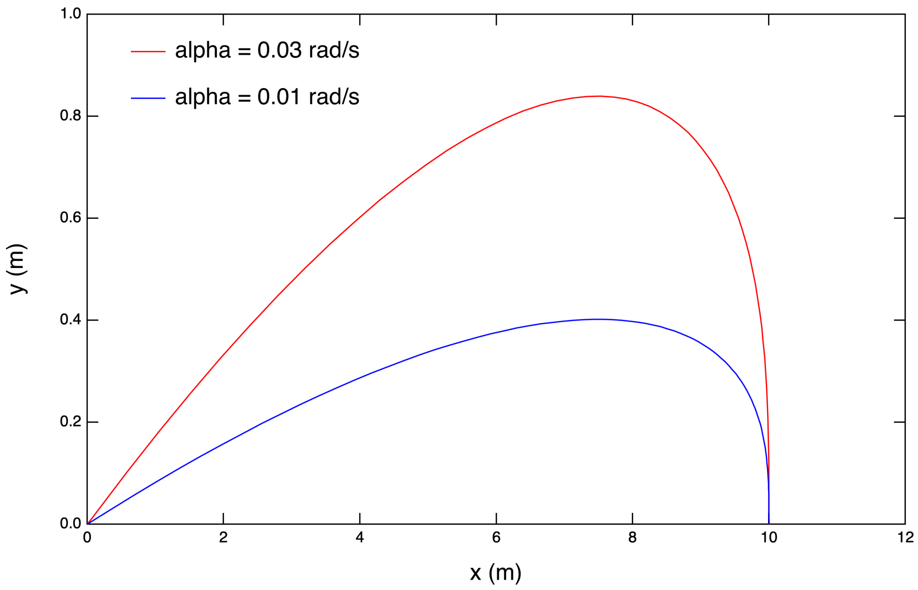


Figure 7.4: Solution of Problem 7.12 with $r_0 = 10$ m and $\alpha = 0.01$ rad/s (blue) and $\alpha = 0.03$ rad/s (red)

$$L = T - U = \frac{1}{2} m \sum_{i=1}^{3} \dot{x}_i^2 - U(x_i) \tag{7.63}$$

The Lagrange equation for this Lagrangian is given by

$$\frac{\partial L}{\partial x_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}_i} = -\frac{\partial U}{\partial x_i} - \frac{d}{dt}\frac{\partial T}{\partial \dot{x}_i} = 0 \tag{7.64}$$

and we can rewrite this as

$$\frac{d}{dt}\frac{\partial T}{\partial \dot{x}_i} = -\frac{\partial U}{\partial x_i} = F_i = m\ddot{x}_i = \frac{d}{dt} p_i \tag{7.65}$$

This last equation suggests that we define the generalized momentum of a particle in the following way:

$$p_i = \frac{\partial T}{\partial \dot{x}_i} \tag{7.66}$$

It is obviously consistent with our definition of linear momentum in Cartesian coordinates.
Consider a particle moving in a two-dimensional plane and having its motion described in terms of spherical coordinates. The kinetic energy of the particle is equal to

$$T = \frac{1}{2} m \left(\dot{r}^2 + r^2 \dot{\theta}^2 \right) \tag{7.67}$$

Since there are two generalized coordinates we can determine two generalized momenta:

$$p_r = \frac{\partial T}{\partial \dot{r}} = m\dot{r} \tag{7.68}$$

which is the linear momentum of the particle, and

$$p_\theta = \frac{\partial T}{\partial \dot{\theta}} = mr^2\dot{\theta} \tag{7.69}$$

which is the angular momentum of the particle. We thus see that two distinct concepts from our introductory courses emerge directly from our Lagrangian theory.

7.6 Homogeneous functions

Consider a homogeneous quadratic function f that depends only on the products of the generalized velocities:

$$f = \sum_{j,k} a_{j,k} \dot{q}_j \dot{q}_k \tag{7.70}$$

An example of such function would be the kinetic energy of a particle. Consider what happens when we differentiate this function with respect to one of the generalized velocities:

$$\frac{\partial f}{\partial \dot{q}_l} = \sum_j a_{j,l} \dot{q}_j + \sum_k a_{l,k} \dot{q}_k \tag{7.71}$$

If we multiply this equation by dq_l/dt and sum over all values of l we obtain:

$$\sum_l \dot{q}_l \frac{\partial f}{\partial \dot{q}_l} = \sum_l \dot{q}_l \left(\sum_j a_{j,l} \dot{q}_j + \sum_k a_{l,k} \dot{q}_k \right) = \sum_{l,j} a_{j,l} \dot{q}_j \dot{q}_l + \sum_{l,k} a_{l,k} \dot{q}_k \dot{q}_l = 2 \sum_{j,k} a_{j,k} \dot{q}_j \dot{q}_k = 2f \tag{7.72}$$

In general, if f is a homogeneous function of the parameter ${y_k}^n$, then

$$\sum_k \dot{q}_k \frac{\partial f}{\partial \dot{q}_k} = nf \tag{7.73}$$

7.7 Conservation of Energy

If we consider a closed system, a system that does not interact with its environment, then we expect that the Lagrangian that describes this system does not depend explicitly on time. That is

$$\frac{\partial L}{\partial t} = 0 \tag{7.74}$$

Of course, this does not mean that $dL/dt = 0$ since

$$\frac{dL}{dt} = \sum_j \frac{\partial L}{\partial q_j} \dot{q}_j + \sum_k \frac{\partial L}{\partial \dot{q}_k} \ddot{q}_j + \frac{\partial L}{\partial t} = \sum_j \frac{\partial L}{\partial q_j} \dot{q}_j + \sum_k \frac{\partial L}{\partial \dot{q}_k} \ddot{q}_j \tag{7.75}$$

Using the Lagrange equations shown in Eq. 7.3, we can rewrite this equation as

$$\frac{dL}{dt} = \sum_j \dot{q}_j \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j} \right) + \sum_k \frac{\partial L}{\partial \dot{q}_k} \ddot{q}_j = \sum_j \frac{d}{dt}\left(\dot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right) \tag{7.76}$$

This equation can be written as

$$\frac{d}{dt}\left(L - \sum_j \left(\dot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right) \right) = 0 \tag{7.77}$$

or

$$L - \sum_j \left(\dot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right) = \text{constant} = -H \tag{7.78}$$

The constant H is called the Hamiltonian of the system and the Hamiltonian is defined as

$$H = \sum_j \left(\dot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right) - L \tag{7.79}$$

The Hamiltonian H is a conserved quantity for the system we are currently considering. Since T is a homogeneous function of order 2, we know that

$$\sum_j \left(\dot{q}_j \frac{\partial T}{\partial \dot{q}_j} \right) = 2T \tag{7.80}$$

and thus

$$\sum_j \left(\dot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right) = \sum_j \left(\dot{q}_j \frac{\partial (T-U)}{\partial \dot{q}_j} \right) = \sum_j \left(\dot{q}_j \frac{\partial T}{\partial \dot{q}_j} \right) = 2T \tag{7.81}$$

Using this relation, we can rewrite Eq. 7.79 as

$$H = \sum_j \left(\dot{q}_j \frac{\partial L}{\partial \dot{q}_j} \right) - L = 2T - (T-U) = (T+U) = E \tag{7.82}$$

In this case, we find that the Hamiltonian of the system is equal to the total energy of the system, and we thus conclude that the total energy is conserved. The equality of H and E is only satisfied if the following conditions are met:

1. The potential U depends only on position, and not on velocity.
2. The transformation rules connecting Cartesian and generalized coordinates are independent of time.

The latter condition is not met in for example a moving coordinate system. In such a system, the Hamiltonian will not be equal to the total energy. We thus conclude that if the Lagrangian of a system does not depend explicitly on time, the total energy of that system will be conserved.

7.7.1 Example Problem 7.22

A particle of mass m moves in one dimension under the influence of a force F:

$$F(x,t) = \frac{k}{x^2} e^{-(t/\tau)} \tag{7.83}$$

where k and τ are positive constants. Compute the Lagrangian and Hamiltonian functions. Compare the Hamiltonian and the total energy and discuss the conservation of energy for the system.

The potential energy U corresponding to this force F must satisfy the relation

$$F = -\frac{\partial U}{\partial x} \tag{7.84}$$

and U must thus be equal to

$$U = \frac{k}{x} e^{-t/\tau} \tag{7.85}$$

Therefore, the Lagrangian is

$$L = T - U = \frac{1}{2} m \dot{x}^2 - \frac{k}{x} e^{-t/\tau} \tag{7.86}$$

The Hamiltonian is given by

$$H = p_x \dot{x} - L = \dot{x} \frac{\partial L}{\partial \dot{x}} - L = m\dot{x}^2 - \left(\frac{1}{2} m \dot{x}^2 - \frac{k}{x} e^{-t/\tau}\right) = \frac{1}{2} m \dot{x}^2 + \frac{k}{x} e^{-t/\tau} \tag{7.87}$$

so that

$$H = \frac{p_x^2}{2m} + \frac{k}{x} e^{-t/\tau} \tag{7.88}$$

The Hamiltonian is equal to the total energy, $T + U$, because the potential does not depend on velocity, but the total energy of the system is not conserved because H contains the time explicitly.

7.8 Conservation of Linear Momentum

The Lagrangian should be unaffected by a translation of the entire system in space, assuming that space is homogeneous (which is one of the requirement of an inertial reference frame). Consider what happens when we carry out an infinitesimal displacement of the coordinate system along one of the coordinate axes. The change in the Lagrangian as a result of this displacement must be equal to zero:

$$\delta L = \frac{\partial L}{\partial x_i}\delta x_i + \frac{\partial L}{\partial \dot{x}_i}\delta \dot{x}_i = 0 \tag{7.89}$$

We can rewrite this equation as

$$\delta L = \frac{\partial L}{\partial x_i}\delta x_i + \frac{\partial L}{\partial \dot{x}_i}\delta \dot{x}_i = \frac{\partial L}{\partial x_i}\delta x_i + \frac{\partial L}{\partial \dot{x}_i}\left(\frac{d}{dt}\delta x_i\right) = \frac{\partial L}{\partial x_i}\delta x_i = 0 \tag{7.90}$$

Since the displacement is arbitrary, this equation can only be correct if

$$\frac{\partial L}{\partial x_i} = 0 \tag{7.91}$$

Using the Lagrange equation this is equivalent to requiring

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}_i} = 0 \tag{7.92}$$

or

$$\frac{\partial L}{\partial \dot{x}_i} = \text{constant} \tag{7.93}$$

Assuming that the potential U does not depend on velocity we see that this relation is equivalent to

$$\frac{\partial L}{\partial \dot{x}_i} = \frac{\partial (T-U)}{\partial \dot{x}_i} = \frac{\partial T}{\partial \dot{x}_i}\frac{\partial}{\partial \dot{x}_i}\left(\frac{(m\dot{x}_i)^2}{2m}\right) = \frac{2m^2\dot{x}_i}{2m} = m\dot{x}_i = p_i = \text{constant} \tag{7.94}$$

The consequence of the independence of the Lagrangian under a translation of space is that linear momentum is conserved.

7.9 Conservation of Angular Momentum

Space is an inertial reference frame is isotropic, which means that the properties of a system are unaffected by the orientation of the system. In this case we expect that the Lagrangian does not change when the coordinate axes are rotated through an infinitesimal angle. A rotation through such an angle produces the following change in the position vector:

$$\delta\overline{r} = \delta\overline{\theta} \times \overline{r} \tag{7.95}$$

The velocity vector will change in the same way:

$$\delta\dot{\overline{r}} = \delta\overline{\theta} \times \dot{\overline{r}} \tag{7.96}$$

The Lagrangian should not change as a result of such a transformation. Thus we must require that

$$\delta L = \sum_i \left(\frac{\partial L}{\partial x_i} \delta x_i + \frac{\partial L}{\partial \dot{x}_i} \delta \dot{x}_i \right) = \sum_i \left\{ \left(\frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} \right) \delta x_i + \frac{\partial L}{\partial \dot{x}_i} \delta \dot{x}_i \right\} = \sum_i \left(\dot{p}_i \delta x_i + p_i \delta \dot{x}_i \right) = 0 \tag{7.97}$$

We thus conclude that

$$\dot{\overline{p}} \cdot \delta \overline{r} + \overline{p} \cdot \delta \dot{\overline{r}} = 0 \tag{7.98}$$

When we express the changes in terms of the rotation angle we obtain:

$$\dot{\overline{p}} \cdot \left(\delta \overline{\theta} \times \overline{r} \right) + \overline{p} \cdot \left(\delta \overline{\theta} \times \dot{\overline{r}} \right) = \delta \overline{\theta} \cdot \left[\dot{\overline{r}} \times \overline{p} + \overline{r} \times \dot{\overline{p}} \right] = \delta \overline{\theta} \cdot \left[\frac{d}{dt} \left(\overline{r} \times \overline{p} \right) \right] = 0 \tag{7.99}$$

Since the angle of rotation was an arbitrary angle, this relation can only be satisfied if

$$\frac{d}{dt} \left(\overline{r} \times \overline{p} \right) = 0 \tag{7.100}$$

or

$$\overline{r} \times \overline{p} = \overline{L} = \text{constant} \tag{7.101}$$

The angular momentum of the system is thus conserved. This conserved quantity is a direct consequence of the invariance of the Lagrangian for infinitesimal rotations. We conclude that the important conserved quantities are a direct consequence of the properties of space (and its symmetries).

7.10 Canonical Equations of Motion

The Lagrangian we have discussed in this Chapter is a function of the generalized position and the generalized velocity. The equations of motion can also be expressed in terms of the generalized position and the generalized momentum. The generalized momentum is defined as

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{7.102}$$

We can use the generalized momentum to rewrite the Lagrange equations of motion:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = \dot{p}_i = \frac{\partial L}{\partial q_i} \tag{7.103}$$

The Hamiltonian can also be expressed in terms of the generalized momentum

$$H = \sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} - L = \sum_j \dot{q}_j p_j - L \tag{7.104}$$

In general, we will write the Hamiltonian in terms of the generalized position and the generalized momentum . The change in H due to small changes in time and in the generalized position and momentum is equal to

$$\begin{aligned} dH &= \textstyle\sum_j \left(\dot{q}_j dp_j + p_j d\dot{q}_j - \frac{\partial L}{\partial q_j} dq_j - \frac{\partial L}{\partial \dot{q}_j} d\dot{q}_j \right) - \frac{\partial L}{\partial t} dt = \\ &= \textstyle\sum_j \left(\dot{q}_j dp_j + p_j d\dot{q}_j - \frac{\partial L}{\partial q_j} dq_j - p_j d\dot{q}_j \right) - \frac{\partial L}{\partial t} dt = \\ &= \textstyle\sum_j \left(\dot{q}_j dp_j - \frac{\partial L}{\partial q_j} dq_j \right) - \frac{\partial L}{\partial t} dt = \\ &= \textstyle\sum_j \left(\dot{q}_j dp_j - \dot{p}_j dq_j \right) - \frac{\partial L}{\partial t} dt \end{aligned} \tag{7.105}$$

The change in H can also be expressed in the following way:

$$dH = \sum_j \left(\frac{\partial H}{\partial p_j} dp_j + \frac{\partial H}{\partial q_j} dq_j \right) + \frac{\partial H}{\partial t} dt \tag{7.106}$$

After combining the last two equations we obtain the following relation:

$$\sum_j \left(\frac{\partial H}{\partial p_j} dp_j + \frac{\partial H}{\partial q_j} dq_j \right) + \frac{\partial H}{\partial t} dt = \sum_j \left(\dot{q}_j dp_j - \dot{p}_j dq_j \right) - \frac{\partial L}{\partial t} dt \tag{7.107}$$

or

$$\sum_j \left(\left\{ \frac{\partial H}{\partial p_j} - \dot{q}_j \right\} dp_j + \left\{ \frac{\partial H}{\partial q_j} + \dot{p}_j \right\} dq_j \right) + \left\{ \frac{\partial H}{\partial t} + \frac{\partial L}{\partial t} \right\} dt = 0 \tag{7.108}$$

Since the variations in time and the generalized position and momenta are independent, the coefficients of dq_i, dp_i, and dt must be zero. Thus:

$$\frac{\partial H}{\partial p_j} - \dot{q}_j = 0 \tag{7.109}$$

$$\frac{\partial H}{\partial q_j} + \dot{p}_j = 0 \tag{7.110}$$

$$\frac{\partial H}{\partial t} + \frac{\partial L}{\partial t} = 0 \tag{7.111}$$

The first two equations are called **Hamilton's equations of motion** or the **canonical equations of motion**. Note:

1. For each generalized coordinate, there are two canonical equations of motion.
2. For each generalized coordinate, there is one Lagrange equation of motion.
3. Each canonical equation of motion is a first-order differential equation.
4. Each Lagrange equation of motion is a second-order differential equation.

Although first-order differential equations are in general easier to solve than second-order differential equations, the Hamiltonian is often more difficult to construct than the Lagrangian since we must express the Hamiltonian in terms of the generalized position and the generalized momentum.

7.10.1 Example: Problem 7.38

The potential for an anharmonic oscillator is $U = kx^2/2 + bx^4/4$ where k and b are constants. Find Hamilton's equations of motion.

The Hamiltonian of the system is

$$H = T + U = \frac{1}{2} m \left(\frac{dx}{dt} \right)^2 + \frac{kx^2}{2} + \frac{bx^4}{4} = \frac{p^2}{2m} + \frac{kx^2}{2} + \frac{bx^4}{4} \tag{7.112}$$

The Hamiltonian motion equations that follow this Hamiltonian are

$$\frac{dx}{dt} = \frac{\partial H}{\partial p} = \frac{p}{m} \tag{7.113}$$

$$\frac{dp}{dt} = -\frac{\partial H}{\partial x} = -(kx + bx^3) \tag{7.114}$$

7.10.2 Example: Problem 7.28.

A particle of mass m is attached to a force center with a force of magnitude k/r^2. Use polar coordinates and find Hamilton's equations of motion.

The force F that is provided fixed the potential U:

$$U = -\frac{k}{r} \tag{7.115}$$

The Lagrangian, expressed in polar coordinates, is thus equal to

$$L = T - U = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) + \frac{k}{r} \tag{7.116}$$

In order to use Hamilton's equations of motion we must express the Hamiltonian in terms of the generalized position and momentum. The following relations can be used to do this:

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r} \Rightarrow \dot{r} = \frac{p_r}{m} \tag{7.117}$$

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta} \Rightarrow \dot{\theta} = \frac{p_\theta}{mr^2} \tag{7.118}$$

Since the coordinate transformations are independent of t, and the potential energy is velocity-independent, the Hamiltonian is the total energy.

$$\begin{aligned} H &= T + U = \tfrac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - \tfrac{k}{r} = \\ &= \tfrac{1}{2}m\left[\tfrac{p_r^2}{m^2} + r^2\tfrac{p_\theta^2}{m^2r^4}\right] - \tfrac{k}{r} = \tfrac{p_r^2}{2m} + \tfrac{p_\theta^2}{2mr^2} - \tfrac{k}{r} \end{aligned} \tag{7.119}$$

Hamilton's equations of motion can now be found easily

$$\dot{r} = \frac{\partial H}{\partial p_r} = \frac{p_r}{m} \tag{7.120}$$

$$\dot{\theta} = \frac{\partial H}{\partial p_\theta} = \frac{p_\theta}{mr^2} \tag{7.121}$$

$$\dot{p}_r = -\frac{\partial H}{\partial r} = \frac{p_\theta^2}{mr^3} - \frac{k}{r^2} \tag{7.122}$$

$$\dot{p}_\theta = -\frac{\partial H}{\partial \theta} = 0 \tag{7.123}$$

7.10.3 Example: Problem 7.24.

Consider a simple plane pendulum consisting of a mass m attached to a string of length ℓ. After the pendulum is set into motion, the length of the string is shortened at a constant rate:

$$\frac{d\ell}{dt} = -\alpha = \text{constant} \tag{7.124}$$

The suspension point remains fixed. Compute the Lagrangian and Hamiltonian functions. Compare the Hamiltonian and the total energy, and discuss the conservation of energy for the system.

The kinetic energy and the potential energy of the system are expressed as

$$T = \frac{1}{2}m\left(\dot{\ell}^2 + \ell^2\dot{\theta}^2\right) = \frac{1}{2}m\left(\alpha^2 + \ell^2\dot{\theta}^2\right) \tag{7.125}$$

$$U = -mg\ell\cos\theta \tag{7.126}$$

The Lagrangian is equal to

$$L = T - U = \frac{1}{2}m\left(\alpha^2 + \ell^2\dot{\theta}^2\right) + mg\ell\cos\theta \tag{7.127}$$

The Hamiltonian is

$$H = p_\theta\dot{\theta} - L = \frac{\partial L}{\partial\dot{\theta}}\dot{\theta} - L = \frac{p_\theta^2}{2m\ell^2} - \frac{1}{2}m\alpha^2 - mg\ell\cos\theta \tag{7.128}$$

which is different from the total energy, $T + U$. The total energy is thus not conserved in this system because work is done on the system and we have

$$\frac{d}{dt}(T + U) \neq 0 \tag{7.129}$$

NOTE: WE WILL SKIP SECTIONS 7.12 AND 7.13 IN THE TEXTBOOK.

8 Central-Force Motion

In this Chapter, we will use the theory we have discussed in Chapter 6 and 7 and apply it to very important problems in physics, in which we study the motion of two-body systems on which central force are acting. We will encounter important examples from astronomy and from nuclear physics.

8.1 Two-Body Systems with a Central Force

Consider the motion of two objects that are effected by a force acting along the line connecting the centers of the objects. To specify the state of the system, we must specify six coordinates (for example, the (x, y, z) coordinates of their centers). The Lagrangian for this system is given by

$$L = \frac{1}{2}m_1\left|\dot{\vec{r}_1}\right|^2 + \frac{1}{2}m_2\left|\dot{\vec{r}_2}\right|^2 - U\left(\vec{r}_1 - \vec{r}_2\right) \tag{8.1}$$

Note: here we have assumed that the potential depends on the position vector between the two objects. This is not the only way to describe the system; we can for example also specify the position of the center-of-mass, R, and the three components of the relative position vector r. In this case, we choose a coordinate system such that the center-of-mass is at rest, and located at the origin. This requires that

$$\vec{R} = \frac{m_1}{m_1 + m_2}\vec{r}_1 + \frac{m_2}{m_1 + m_2}\vec{r}_2 = 0 \tag{8.2}$$

The relative position vector is defined as

$$\vec{r} = \vec{r}_1 - \vec{r}_2 \tag{8.3}$$

The position vectors of the two masses can be expressed in terms of the relative position vector:

$$\vec{r}_1 = \frac{m_2}{m_1 + m_2}\vec{r} \tag{8.4}$$

$$\vec{r}_2 = \frac{m_1}{m_1 + m_2}\vec{r} \tag{8.5}$$

The Lagrangian can now be rewritten as

$$L = \frac{1}{2}m_1\left(\frac{m_2}{m_1 + m_2}\right)^2\left|\dot{\vec{r}}\right|^2 + \frac{1}{2}m_2\left(\frac{m_1}{m_1 + m_2}\right)^2\left|\dot{\vec{r}}\right|^2 - U\left(\vec{r}\right) = \frac{1}{2}\mu\left|\dot{\vec{r}}\right|^2 - U\left(\vec{r}\right) \tag{8.6}$$

where μ is the reduced mass of the system:

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \tag{8.7}$$

8.2 Two-Body Systems with a Central Force: Conserved Quantities

Since we have assumed that the potential U depends only on the relative position between the two objects, the system poses spherical symmetry. As we have seen in Chapter 7, this type of symmetry implies that the angular momentum of the system is conserved. As a result, the momentum and position vector will lay in a plane, perpendicular to the angular momentum vector, which is fixed in space. The three-dimensional problem is thus reduced to a two-dimensional problem. We can express the Lagrangian in terms of the radial distance r and the polar angle θ:

$$L = \frac{1}{2}\mu\left(\dot{r}^2 + r^2\dot{\theta}^2\right) - U\left(r\right) \tag{8.8}$$

The generalized momenta for this Lagrangian are

$$p_r = \frac{\partial L}{\partial \dot{r}} = \mu\dot{r} \tag{8.9}$$

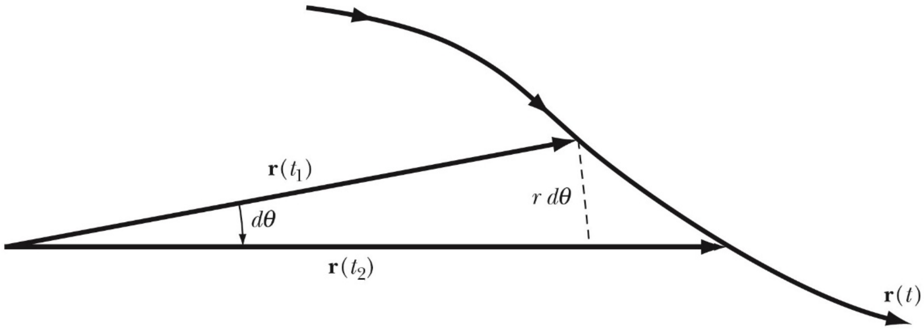


Figure 8.1: Calculation of the areal velocity.

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = \mu r^2 \dot{\theta} \tag{8.10}$$

The Lagrange equations can be used to determine the derivative of these momenta with respect to time:

$$\dot{p}_r = \frac{d}{dt}\frac{\partial L}{\partial \dot{r}} = \frac{\partial L}{\partial r} = \mu r \dot{\theta}^2 - \frac{\partial U}{\partial r} \tag{8.11}$$

$$\dot{p}_\theta = \frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} = \frac{\partial L}{\partial \theta} = 0 \tag{8.12}$$

The last equation tells us that the generalized momentum p_q is constant:

$$\ell = \mu r^2 \dot{\theta} = \text{constant} \tag{8.13}$$

The constant ℓ is related to the areal velocity. Consider the situation in Fig. 8.1. During the time interval dt, the radius vector sweeps an area dA where

$$dA = \frac{1}{2} r^2 d\theta \tag{8.14}$$

The areal velocity, dA/dt, is thus equal to

$$\frac{dA}{dt} = \frac{1}{2} r^2 \frac{d\theta}{dt} = \frac{1}{2} r^2 \frac{\ell}{\mu r^2} = \frac{\ell}{2\mu} = \text{constant} \tag{8.15}$$

This result is also known as **Kepler's Second Law**.
The Lagrangian for the two-body system does not depend explicitly on time. In Chapter 7 we showed that in that case, the energy of the system is conserved. The total energy E of the system is equal to

$$E = T + U = \frac{1}{2}\mu\left(\dot{r}^2 + r^2\dot{\theta}^2\right) + U(r) = \frac{1}{2}\mu\left(\dot{r}^2 + r^2\left(\frac{\ell}{\mu r^2}\right)^2\right) + U(r) = \frac{1}{2}\mu\dot{r}^2 + \frac{1}{2}\frac{l^2}{\mu r^2} + U(r) \tag{8.16}$$

This equation can be rewritten as

$$\dot{r}^2 = \frac{2}{\mu}(E - U(r)) - \frac{\ell^2}{\mu^2 r^2} \tag{8.17}$$

8.3 Two-Body Systems with a Central Force: Equations of Motion

If the potential energy is specified, we can use the expression for the total energy E to determine dr/dt:

$$\frac{dr}{dt} = \pm\sqrt{\frac{2}{\mu}(E-U) - \frac{l^2}{\mu^2 r^2}} \tag{8.18}$$

This equation can be used to find the time t as function of r:

$$t = \int dt = \pm\int \frac{1}{\sqrt{\frac{2}{\mu}(E-U(r)) - \frac{l^2}{\mu^2 r^2}}}dr \tag{8.19}$$

However, in many cases, the shape of the trajectory, $\theta(r)$, is more important than the time dependence. We can express the change in the polar angle in terms of the change in the radial distance:

$$d\theta = \frac{d\theta}{dt}\frac{dt}{dr}dr = \frac{\dot{\theta}}{\dot{r}}dr \tag{8.20}$$

Integrating both sides we obtain the following orbital equation

$$\theta(r) = \int \frac{\dot{\theta}}{\dot{r}}dr = \pm\int \frac{\frac{\ell}{\mu r^2}}{\sqrt{\frac{2}{\mu}(E-U) - \frac{l^2}{\mu^2 r^2}}}dr = \pm\int \frac{\ell}{r^2\sqrt{2\mu\left(E-U-\frac{l^2}{2\mu r^2}\right)}}dr \tag{8.21}$$

The extremes of the orbit can be found in general by requiring that $dr/dt = 0$, or

$$\frac{2}{\mu}(E-U) - \frac{l^2}{\mu^2 r^2} = \frac{2}{\mu}\left(E - U(r) - \frac{l^2}{2\mu r^2}\right) = 0 \tag{8.22}$$

In general, this equation has two solutions, and the orbit is confined between a minimum and maximum value of r. Under certain conditions, there is only a single solution, and in that case the orbit is circular. Using the orbital equation we can determine the change in the polar angle when the radius changes from r_{min} to r_{max}. During one period, the polar angle will change by

$$\Delta\theta = 2\int_{r_{\min}}^{r_{\max}} \frac{\ell}{r^2\sqrt{2\mu\left(E-U-\frac{l^2}{2\mu r^2}\right)}}dr \tag{8.23}$$

If the change in the polar angle is a rational fraction of 2π then after a number of complete orbits, the system will have returned to its original position. In this case, the orbit is closed. In all other cases, the orbit is open.

The orbital motion is specified above in terms of the potential U. Another approach to study the equations of motion is to start from the Lagrange equations. In this case we obtain an equation of motion that includes the force F instead of the potential U:

$$\frac{d^2}{d\theta^2}\left(\frac{1}{r}\right) + \frac{1}{r} = -\frac{\mu r^2}{l^2}F(r) \tag{8.24}$$

This version of the equations of motion is useful when we can measure the orbit and want to find the force that produces this orbit.

8.3.1 Example: Problem 8.8

Investigate the motion of a particle repelled by a force center according to the law $F(r) = kr$. Show that the orbit can only be hyperbolic.

The general expression for $\theta(r)$ is [see Eq. (8.17) in the text book]

$$\theta(r) = \int \frac{\left(\ell/r^2\right)dr}{\sqrt{2\mu\left[E - U - \frac{\ell^2}{2\mu r^2}\right]}} \tag{8.25}$$

where

$$U = -\int krdr = -kr^2/2 \tag{8.26}$$

in the present case. Substituting $x = r^2$ and $dx = 2rdr$ into Eq. (8.25), we have

$$\theta(r) = \frac{1}{2}\int \frac{dx}{x\sqrt{\frac{\mu k}{\ell^2}x^2 + \frac{2\mu E}{\ell^2}x - 1}} \tag{8.27}$$

Using Eq. (E.10b), Appendix E,

$$\int \frac{dx}{x\sqrt{ax^2 + bx + c}} = \frac{1}{\sqrt{-c}}\sin^{-1}\left[\frac{bx + 2c}{|x|\sqrt{b^2 - 4ac}}\right] \tag{8.28}$$

and expressing again in terms of r, we find

$$\theta(r) = \frac{1}{2}\sin^{-1}\left[\frac{\left[\frac{\mu E}{\ell^2}r^2 - 1\right]}{r^2\sqrt{\frac{\mu^2 E^2}{\ell^4} + \frac{\mu k}{\ell^2}}}\right] + \theta_0 \tag{8.29}$$

or,

$$\sin 2(\theta - \theta_0) = \frac{1}{\sqrt{1 + \frac{\ell^2 k}{\mu E^2}}} - \frac{1}{r^2}\frac{\ell^2/\mu E}{\sqrt{1 + \frac{\ell^2 k}{\mu E^2}}} \tag{8.30}$$

In order to interpret this result, we define

$$\sqrt{1 + \frac{\ell^2 k}{\mu E^2}} \equiv \varepsilon' \tag{8.31}$$

$$\frac{\ell^2}{\mu E} \equiv \alpha' \tag{8.32}$$

and specifying $\theta_0 = \pi/4$, Eq. (8.30) becomes

$$\frac{\alpha'}{r^2} = 1 + \varepsilon'\cos 2\theta \tag{8.33}$$

or,

$$\alpha' = r^2 + \varepsilon' r^2\left(\cos^2\theta - \sin^2\theta\right) \tag{8.34}$$

Rewriting Eq. (8.34) in x-y coordinates, we find

$$\alpha' = x^2 + y^2 + \varepsilon'\left(x^2 - y^2\right) \tag{8.35}$$

or,

$$1 = \frac{x^2}{\frac{\alpha'}{1+\varepsilon'}} + \frac{y^2}{\frac{\alpha'}{1-\varepsilon'}} \tag{8.36}$$

Since $a' > 0$, $\varepsilon > 1$ from the definition, Eq. (8.36) is equivalent to

$$1 = \frac{x^2}{\frac{\alpha'}{1+\varepsilon'}} - \frac{y^2}{\frac{\alpha'}{|1-\varepsilon'|}} \tag{8.37}$$

which is the equation of a hyperbola.

8.4 Solving the Orbital Equation

The orbital equation can only be solved analytically for certain force laws. Consider for example the gravitational force. The corresponding potential is $-k/r$ and the polar angle θ is thus equal to

$$\theta(r) = \pm \int \frac{\ell/r^2}{\sqrt{2\mu\left(E + \frac{k}{r} - \frac{\ell^2}{2\mu r^2}\right)}} dr \tag{8.38}$$

Consider the change of variables from r to $u = \ell/r$:

$$\begin{aligned}
\theta(r) &= \pm \int \frac{(\ell/r)^2/\ell}{\sqrt{2\mu\left(E + \frac{k}{\ell}\frac{\ell}{r} - \frac{1}{2\mu}\left(\frac{\ell}{r}\right)^2\right)}} d\left(\frac{\ell}{u}\right) = \\
&= \pm \int \frac{u^2/\ell}{\sqrt{2\mu\left(E + \frac{k}{\ell}u - \frac{1}{2\mu}u^2\right)}} \frac{\ell}{u^2} du = \\
&= \pm \int \frac{1}{\sqrt{2\mu\left(E + \frac{k}{\ell}u - \frac{1}{2\mu}u^2\right)}} du
\end{aligned} \tag{8.39}$$

The integral can be solved using one of the integrals found in Appendix E (see E8.c):

$$\begin{aligned}
\theta(r) &= \pm \int \frac{1}{\sqrt{-u^2 + 2\mu\frac{k}{\ell}u + 2\mu E}} du = \pm \sin^{-1}\left(\frac{-2u + 2\mu\frac{k}{\ell}}{\sqrt{\left(2\mu\frac{k}{\ell}\right)^2 + 8\mu E}}\right) + C = \\
&= \pm \sin^{-1}\left(\frac{\mu\frac{k}{\ell} - u}{\sqrt{\left(\mu\frac{k}{\ell}\right)^2 + 2\mu E}}\right) + C = \pm \sin^{-1}\left(\frac{\mu\frac{k}{\ell} - \frac{\ell}{r}}{\sqrt{\left(\mu\frac{k}{\ell}\right)^2 + 2\mu E}}\right) + C = \\
&= \pm \sin^{-1}\left(\frac{\mu k - \frac{\ell^2}{r}}{\sqrt{(\mu k)^2 + 2\mu\ell^2 E}}\right) + C
\end{aligned} \tag{8.40}$$

This equation can be rewritten as

$$\sin(\theta + \text{constant}) = \frac{\mu k - \frac{\ell^2}{r}}{\sqrt{(\mu k)^2 + 2\mu\ell^2 E}} \tag{8.41}$$

We can always choose our reference position such that the constant is equal to $\pi/2$ and we thus find the following solution:

$$\cos(\theta) = \frac{\mu k - \frac{\ell^2}{r}}{\sqrt{(\mu k)^2 + 2\mu\ell^2 E}} \tag{8.42}$$

We can rewrite this expression such that we can determine the distance r as function of the polar angle:

$$r = \frac{\ell^2}{\mu k - \sqrt{(\mu k)^2 + 2\mu\ell^2 E}\cos(\theta)} = \frac{\ell^2}{\mu k\left(1 - \sqrt{1 + \frac{2}{\mu}\frac{\ell^2}{k^2}E}\cos(\theta)\right)} \tag{8.43}$$

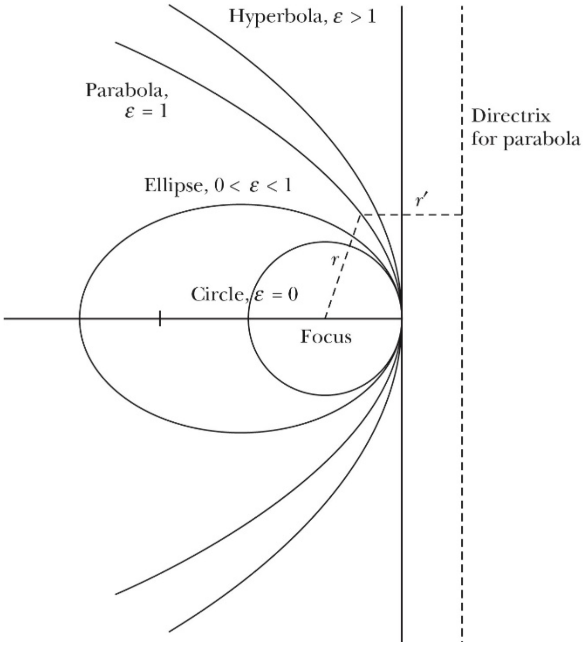


Figure 8.2: Possible orbits in the gravitational field.

Since cosθ varies between -1 and +1, we see that the minimum (the **pericenter**) and the maximum (the **apocenter**) positions are

$$r_{\text{min}} = \frac{\ell^2}{\mu k\left(1 + \sqrt{1 + \frac{2}{\mu}\frac{\ell^2}{k^2}E}\right)} \tag{8.44}$$

$$r_{\text{max}} = \frac{\ell^2}{\mu k\left(1 - \sqrt{1 + \frac{2}{\mu}\frac{\ell^2}{k^2}E}\right)} \tag{8.45}$$

The equation for the orbit is in general expressed in terms of the **eccentricity** ε and the **latus rectum** 2α:

$$\varepsilon = \sqrt{1 + \frac{2}{\mu}\frac{\ell^2}{k^2}E} \tag{8.46}$$

$$\alpha = \frac{l^2}{\mu k} \tag{8.47}$$

The possible orbits are usually parameterized in terms of the eccentricity, and examples are shown in Fig. 8.2.
The period of the orbital motion can be found by integrating the expression for dt over one complete period:

$$\tau = \int dt = \frac{2\mu}{\ell}\int dA = \frac{2\mu}{\ell}(\pi ab) = \frac{2\mu}{\ell}\left(\pi\frac{k}{2|E|}\frac{\ell}{\sqrt{2\mu|E|}}\right) = \pi k\sqrt{\frac{\mu}{2}}|E|^{-3/2} \tag{8.48}$$

When we take the square of this equation we get Kepler's third law:

$$\tau^2 = \pi^2 k^2 \frac{\mu}{2}|E|^{-3} = \pi^2 k^2 \frac{\mu}{2}\left(\frac{2a}{k}\right)^3 = \frac{4\pi^2\mu}{k}a^3 \tag{8.49}$$

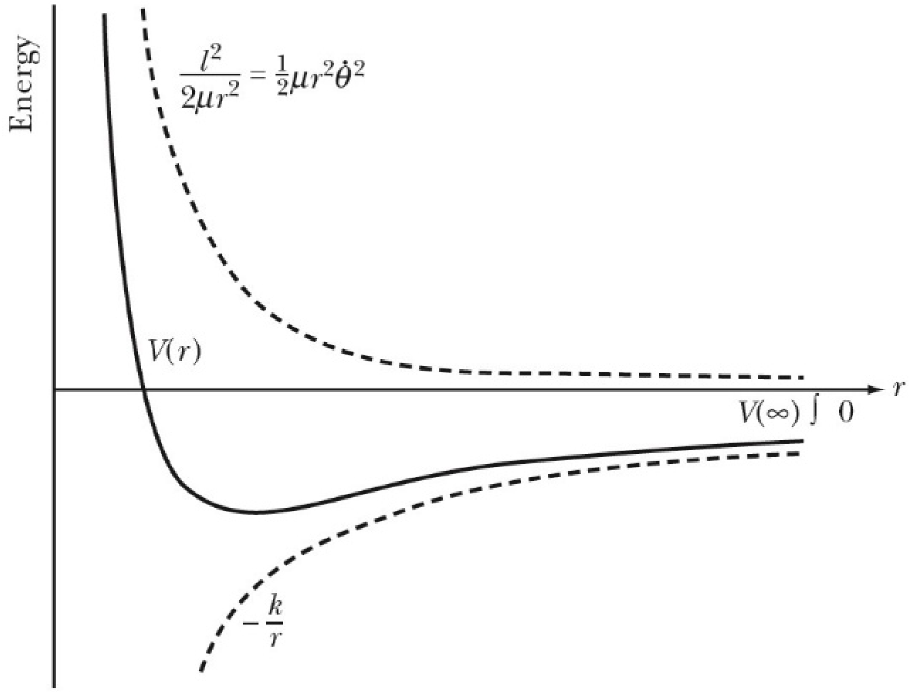


Figure 8.3: The effective potential for the gravitational force when the system has an angular momentum ℓ.

8.5 The Centripetal Force and Potential

In the previous discussion it appears as if the potential U is modified by the term $\ell^2/(2mr^2)$. This term depends only on the position r since ℓ is constant, and it is interpreted as a potential energy. The force associated with this potential energy is

$$F_c = -\frac{\partial U_c}{\partial r} = \frac{\ell^2}{\mu r^3} = \mu r\dot{\theta}^2 = \mu\frac{\left(r\dot{\theta}\right)^2}{r} \tag{8.50}$$

This force is often called the **centripetal force** (although it is not a real force), and the potential is called the **centripetal potential**. This potential is a fictitious potential and it represents the effect of the angular momentum about the origin. Figure 8.3 shows an example of the real potential, due to the gravitational force in this case, and the centripetal potential. The effective potential is the sum of these two potentials and has a characteristic dip where the potential energy has a minimum. The result of this dip is that there are certain energies for which the orbit is bound (has a minimum and maximum distance). These turning points are called the **apsidal distances** of the orbit. We also note that at small distances, the force becomes repulsive.

8.5.1 Example: Problem 8.22

Discuss the motion of a particle moving in an attractive central-force field described by $F(r) = -k/r^3$. Describe some of the orbits for different values of the total energy.

For the given force

$$F(r) = -\frac{k}{r^3} \tag{8.51}$$

the potential is

$$U(r) = -\frac{k}{2r^2} \tag{8.52}$$

and the effective potential is

$$V(r) = \frac{1}{2}\left[\frac{\ell^2}{\mu} - k\right]\frac{1}{r^2} \tag{8.53}$$

The equation of the orbit is [cf. Eq. (8.20) in the text book]

$$\frac{d^2u}{d\theta^2} + u = -\frac{\mu}{\ell^2 u^2}\left(-ku^3\right) \tag{8.54}$$

or

$$\frac{d^2u}{d\theta^2} + \left[1 - \frac{\mu k}{\ell^2}\right]u = 0 \tag{8.55}$$

Let us consider the motion for various values of ℓ.

1. $\ell^2 = \mu k$:

   In this case the effective potential $V(r)$ vanishes and the orbit equation is

   $$\frac{d^2u}{d\theta^2} = 0 \tag{8.56}$$

   with the solution

   $$u = \frac{1}{r} = A\theta + B \tag{8.57}$$

   We conclude that the particle spirals towards the force center.

2. $\ell^2 > \mu k$:

   In this case the effective potential is positive and decreases monotonically with increasing r. For any value of the total energy E, the particle will approach the force center and will undergo a reversal of its motion at $r = r_0$; the particle will then proceed again to an infinite distance. Setting

   $$1 - \mu k/\ell^2 \equiv \beta^2 > 0 \tag{8.58}$$

   Eq. (8.55) becomes

   $$\frac{d^2u}{d\theta^2} + \beta^2 u = 0 \tag{8.59}$$

   with the solution

   $$u = \frac{1}{r} = A\cos(\beta\theta - \delta) \tag{8.60}$$

   Since the minimum value of u is zero, this solution corresponds to unbounded motion, as expected from the form of the effective potential $V(r)$.

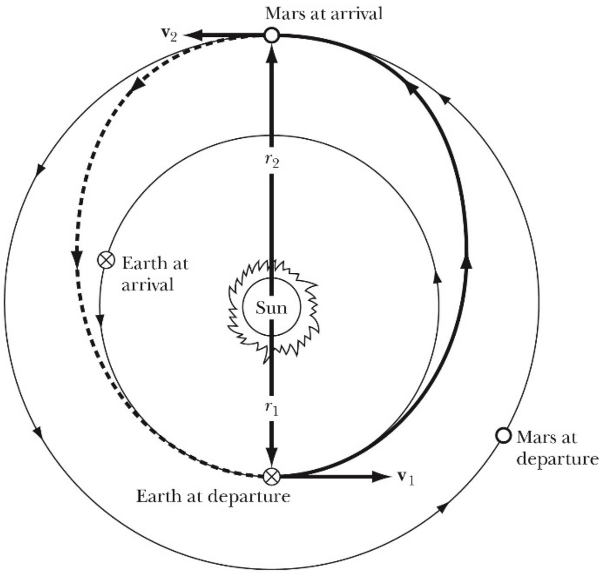


Figure 8.4: The Hohmann transfer to travel from Earth to Mars.

3. $\ell^2 < \mu k$:

   For this case we set

$$\mu k/\ell^2 - 1 \equiv G^2 > 0 \tag{8.61}$$

   and the orbit equation becomes

$$\frac{d^2u}{d\theta^2} - G^2 u = 0 \tag{8.62}$$

   with the solution

$$u = \frac{1}{r} = A\cosh(\beta\theta - \delta) \tag{8.63}$$

   The particle thus spirals in towards the force center.

8.6 Orbital Motion

The understanding of orbital dynamics is very important for space travel. The orbit in which a spaceship travels is determined by the energy of the spaceship. When we change the energy of the ship, we will change the orbit from for example a spherical orbit to an elliptical orbit. By changing the velocity at the appropriate point, we can control the orientation of the new orbit.

The Hohmann transfer represents the path of minimum energy expenditure to move from one solar-based orbit to another. Consider travel from Earth to Mars (see Fig. 8.4). The goal is to get our spaceship in an orbit that has apsidal distances that correspond to the distance between the Earth and the sun and between mars and the sun. This requires that

$$r_1 = a(1-\varepsilon) \tag{8.64}$$

and

$$r_2 = a(1+\varepsilon) \tag{8.65}$$

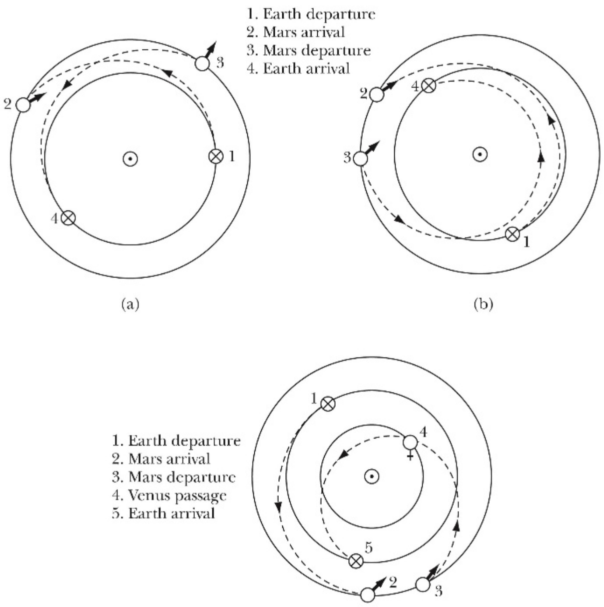


Figure 8.5: Different ways to get from Earth to Mars.

The eccentricity of such an orbit is thus equal to

$$\varepsilon = \frac{r_2 - r_1}{2a} \tag{8.66}$$

The total energy of an orbit with a major axis of $a = (r_1 + r_2)/2$ is equal to

$$E = \frac{k}{2a} = \frac{k}{(r_1 + r_2)} \tag{8.67}$$

Since the space ship starts from a circular orbit with a major axis $a = r_1$, its initial energy is equal to

$$E = -\frac{k}{2r_1} \tag{8.68}$$

The increase in the total energy is thus equal to

$$\Delta E = -\frac{k}{(r_1 + r_2)} - \left(-\frac{k}{2r_1}\right) = \left(\frac{k}{2}\right)\left\{\frac{1}{r_1} - \frac{2}{(r_1 + r_2)}\right\} = \left(\frac{k}{2}\right)\frac{r_2 - r_1}{r_1\,(r_1 + r_2)} \tag{8.69}$$

This energy must be provided by the thrust of the engines that increase the velocity of the space ship (note: the potential energy does not change at the moment of burn, assuming the thrusters are only fired for a short period of time).

The problem with the Hohmann transfer mechanism is that the conditions have to be just right, and only if the planets are in the proper position will the transfer work. There are many other ways to travel between Earth and mars. Many of these require less time than the time required for the Hohmann transfer, but they require more fuel. Examples are shown in Fig. 8.5.

SECTIONS 8.9 AND 8.10 WILL BE SKIPPED!

9 Dynamics of a System of Particles

In this Chapter we expand our discussion from the two-body systems discussed in Chapter 8 to systems that consist out of many particles. In general, these particles are exposed to both external and internal forces. In our discussion in this Chapter, we will make the following assumptions about the internal forces:

1. The forces exerted between any two particles are equal in magnitude and opposite in direction.
2. The forces exerted between any two particles are directed parallel or anti-parallel to the line joining the two particles.

These two requirements are fulfilled for many forces. However, there are important forces, such as the magnetic force, that do not satisfy the second assumption.

9.1 The Center-of-Mass

As we discussed in Chapter 8, it is often useful to separate the motion of a system into the motion of its center of mass and the motion of its component relative to the center of mass. The definition of the position of the center of mass for a multi-particle system, shown in Fig. 9.1, is similar to its definition for a two-body system:

$$\vec{R}_{cm} = \frac{\sum_i m_i \vec{r}_i}{\sum_i m_i} = \frac{1}{M}\sum_i m_i \vec{r}_i \tag{9.1}$$

If the mass distribution is a continuous distribution, the summation must be replaced by an integration:

$$\vec{R}_{cm} = \frac{1}{M}\int \vec{r}\, dm \tag{9.2}$$

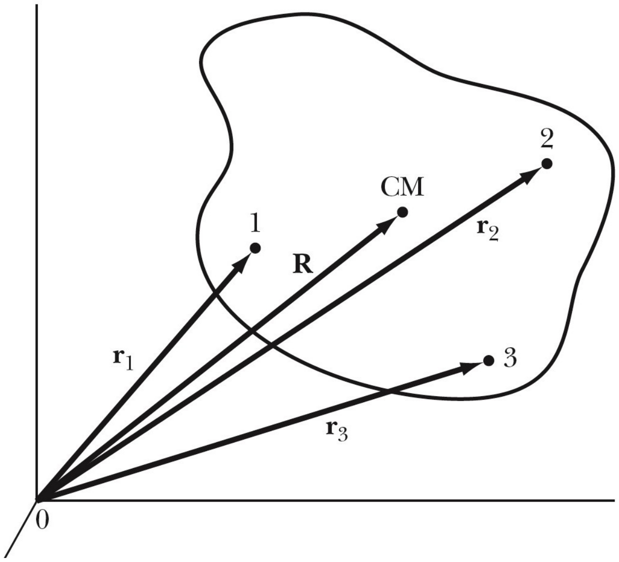


Figure 9.1: The location of the center of mass of a multi-particle system.

9.1.1 Example: Problem 9.1

Find the center of mass of a hemispherical shell of constant density and inner radius r_1 and outer radius r_2.

Put the shell in the $z > 0$ region, with its base in the xy plane. By symmetry,

$$\vec{x}_{cm} = \vec{y}_{cm} = 0 \tag{9.3}$$

To find the z coordinate of the center-of-mass, we divide the shell into thin slices, parallel to the xy plane.

$$\vec{z}_{cm} = \frac{\int \rho z dV}{\int \rho dV} = \frac{\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi/2}\int_{r=r_1}^{r_2} \rho z r^2 \sin\theta dr d\theta d\phi}{\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi/2}\int_{r=r_1}^{r_2} \rho r^2 \sin\theta dr d\theta d\phi} \tag{9.4}$$

Using $z = r\cos\theta$ and evaluating the integrals we obtain

$$\vec{z}_{cm} = \frac{\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi/2}\int_{r=r_1}^{r_2} \rho r^3 \cos\theta \sin\theta dr d\theta d\phi}{\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi/2}\int_{r=r_1}^{r_2} \rho r^2 \sin\theta dr d\theta d\phi} = \frac{\rho(2\pi)\left(\frac{1}{2}\right)\left(\frac{1}{4}\left(r_2^4 - r_1^4\right)\right)}{\rho(2\pi)(1)\left(\frac{1}{3}\left(r_2^3 - r_1^3\right)\right)} = \frac{3\left(r_2^4 - r_1^4\right)}{8\left(r_2^3 - r_1^3\right)} \tag{9.5}$$

9.2 Linear Momentum

Consider a system of particles with a total mass M, exposed to internal and external forces. The linear momentum for this system is defined as

$$\vec{P} = \sum_i m_i \dot{\vec{r}}_i = \frac{d}{dt}\sum_i m_i \vec{r}_i = \frac{d}{dt}\left(M\vec{R}\right) = M\dot{\vec{R}} \tag{9.6}$$

The change in the linear momentum of the system can be expressed in terms of the forces acting on all the particles that make up the system:

$$\frac{d\vec{P}}{dt} = \sum_i m_i \ddot{\vec{r}}_i = \sum_i \vec{F}_i = \sum_i \vec{F}_{i,ext} + \sum_i \sum_{j \neq i} \vec{F}_{ij,\text{int}} = \vec{F}_{ext} \tag{9.7}$$

We see that the linear momentum is constant if the net external force acting on the system is 0 N. If there is an external force acting on the system, the component of the linear momentum in the direction of the net external force is not conserved, but the components in the directions perpendicular to the direction of the net external force are conserved.
We conclude that the linear momentum of the system has the following properties:

1. The center of mass of thids system moves as if it were a single particle with a mass equal to the total mass of the system, M, acted on by the total external force, and independent of the nature of the internal forces.
2. The linear momentum of a system of particles is the same as that of a single particle of mass M, located at the position of the center of mass, and moving in the manner the center of mass is moving.
3. The total linear momentum for a system free of external forces is constant and equal to the linear momentum of the center of mass.

9.3 Angular Momentum

Consider a system of particles that are distributed as shown in Fig. 9.2. We can specify the location of the center of mass of this system by specifying the vector R. This position vector may be time dependent. The location of each component of this system can be specified by either specifying the position vector, r_α, with respect to the origin of the coordinate system, or by specifying the position of the component with respect to the center of mass, r'_α.
The angular momentum of this system with respect to the origin of the coordinate system is equal to

$$\vec{L} = \sum_\alpha \vec{L}_\alpha = \sum_\alpha \left\{\vec{r}_\alpha \times m_\alpha \dot{\vec{r}}_\alpha\right\} = \sum_\alpha \left\{\left(\vec{R} + \vec{r}\,'_\alpha\right) \times m_\alpha \left(\dot{\vec{R}} + \dot{\vec{r}}\,'_\alpha\right)\right\} \tag{9.8}$$

Since

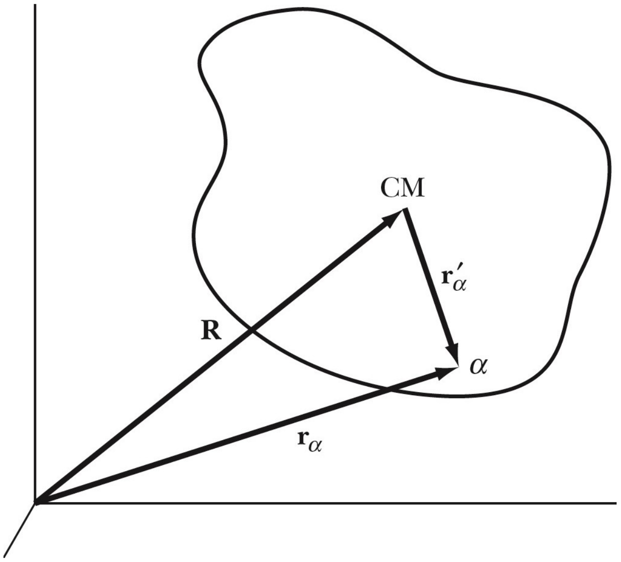


Figure 9.2: Coordinate system used to describe a system of particles.

$$\sum_\alpha \{m_\alpha \vec{r}'_\alpha\} = \sum_\alpha \left\{m_\alpha \left(\vec{r}_\alpha - \vec{R}\right)\right\} = \sum_\alpha \{m_\alpha \vec{r}_\alpha\} - \vec{R}\sum_\alpha m_\alpha = M\vec{R} - M\vec{R} = 0 \tag{9.9}$$

we can rewrite the expression for the angular momentum as

$$\vec{L} = \left(\vec{R} \times \dot{\vec{R}}\right)\sum_\alpha m_\alpha + \sum_\alpha \left\{\vec{r}'_\alpha \times m_\alpha \dot{\vec{r}}'_\alpha\right\} = \vec{R} \times \vec{P} + \sum_\alpha \{\vec{r}'_\alpha \times \vec{p}'_\alpha\} = \vec{L}_{cm} + \vec{L}_{wrt,cm} \tag{9.10}$$

The angular momentum is thus equal to sum of the angular momentum of the center of mass and the angular momentum of the system with respect to the center of mass.
The rate of change of the angular momentum of the system can be determined by using the following relation:

$$\frac{d\vec{L}}{dt} = \dot{L} = \frac{d}{dt}(\vec{r} \times \vec{p}) = \dot{\vec{r}} \times \vec{p} + \vec{r} \times \dot{\vec{p}} = \vec{r} \times \dot{\vec{p}} = \vec{r} \times \vec{F} \tag{9.11}$$

For the system we are currently discussing we can thus conclude that

$$\frac{d\vec{L}}{dt} = \sum_i \left\{\vec{r}_i \times \vec{F}_{i,ext}\right\} + \sum_i \sum_{j \neq i} \left\{\vec{r}_i \times \vec{F}_{ij,\text{int}}\right\} = \sum_i \left\{\vec{r}_i \times \vec{F}_{i,ext}\right\} + \sum_i \sum_{j<i} \left\{\left(\vec{r}_i - \vec{r}_j\right) \times \vec{F}_{ij,\text{int}}\right\} = \sum_i \left\{\vec{r}_i \times \vec{F}_{i,ext}\right\} \tag{9.12}$$

The last step in this derivation is only correct if the internal force between *i* and *j* is parallel or anti-parallel to the relative position vector, but this was one of the two assumptions we made about the internal forces at the beginning of this Chapter. Since the vector product between the position vector and the force vector is the torque *N* associated with this force, we can rewrite the rate of change of the angular momentum of the system as

$$\frac{d\vec{L}}{dt} = \sum_i \vec{N}_{i,ext} \tag{9.13}$$

We conclude that the angular momentum of the system has the following properties:

1. The total angular momentum about an origin is the sum of the angular momentum of the center of mass about that origin and the angular momentum of the system about the position of the center of mass.

2. If the net resultant torques about a given axis vanish, then the total angular momentum of the system about that axis remains constant in time.
3. The total internal torque vanishes if the internal forces are central, and the angular momentum of an isolated system can not be altered without the application of external forces.

9.3.1 Example: Problem 9.13

Even though the total force on a system of particles is zero, the net torque may not be zero. Show that the net torque has the same value in any coordinate system.

The total force acting on the system can be rewritten in terms of the external and internal forces:

$$\vec{F}_{total} = \sum_i \vec{F}_{i,ext} + \sum_i \sum_{j \neq i} \vec{F}_{ij,\text{int}} = \sum_i \vec{F}_{i,ext} \tag{9.14}$$

The problem states that the total force is equal to zero.

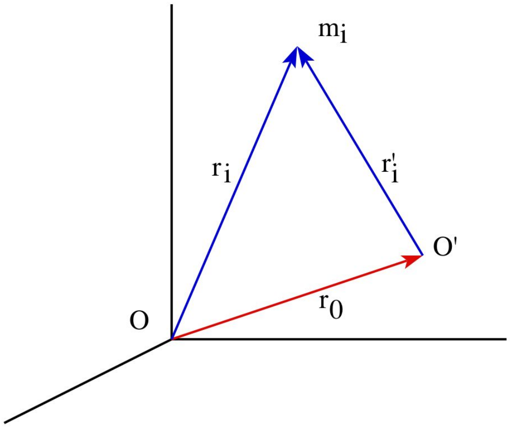


Figure 9.3: Coordinate system used to solve Problem 9.13.

Now consider two coordinate systems, shown in Fig. 9.3, with origins at O and O’ where

1. $\vec{r}_O$ is a vector from O to O’.
2. $\vec{r}_i$ is the position vector of m_i in O.
3. $\vec{r}_i'$ is the position vector of m_i in O’.

We see that $\vec{r}_i = \vec{r}_O + \vec{r}_i'$. The torque with respect to O is given by

$$\tau = \sum_i \vec{r}_i \times \vec{F}_{i,ext} \tag{9.15}$$

The torque with respect to O’ is equal to

$$\begin{aligned}\vec{\tau}' &= \textstyle\sum_i \vec{r}_i' \times \vec{F}_{i,ext} = \\ &= \textstyle\sum_i (\vec{r}_i - \vec{r}_O) \times \vec{F}_{i,ext} \\ &= \textstyle\sum_i \vec{r}_i \times \vec{F}_{i,ext} - \sum_i \vec{r}_O \times \vec{F}_{i,ext} \\ &= \vec{\tau} - \vec{r}_O \times \textstyle\sum_i \vec{F}_{i,ext}\end{aligned} \tag{9.16}$$

But, since it is given that

$$\sum_i \vec{F}_{i,ext} = 0 \tag{9.17}$$

we conclude that

$$\vec{\tau} = \vec{\tau}' \tag{9.18}$$

9.4 Energy

The total energy of a system of particles is equal to the sum of its kinetic and its potential energy.
The kinetic energy of the system is equal to the sum of the kinetic energy of each of the components. The kinetic energy of particle *i* can either be expressed in terms of its velocity with respect to the origin of the coordinate system, or in terms of its velocity with respect to the center of mass:

$$\dot{r}_i^2 = \vec{\dot{r}}_i \cdot \vec{\dot{r}}_i = \left(\vec{\dot{R}} + \vec{\dot{r}}_i{}'\right) \cdot \left(\vec{\dot{R}} + \vec{\dot{r}}_i{}'\right) = \vec{\dot{R}} \cdot \vec{\dot{R}} + 2\vec{\dot{r}}_i{}' \cdot \vec{\dot{R}} + \vec{\dot{r}}_i{}' \cdot \vec{\dot{r}}_i{}' \tag{9.19}$$

The kinetic energy of the system is thus equal to

$$T = \sum_i \frac{1}{2} m_i {v_i}^2 = \sum_i \frac{1}{2} m_i {v_i}'^2 + \sum_i m_i \left(\vec{v}_i{}' \cdot \vec{V}\right) + \sum_i \frac{1}{2} m_i V^2 = \sum_i \frac{1}{2} m_i {v_i}'^2 + \vec{V} \cdot \frac{d}{dt} \sum_i m_i \vec{r}_i{}' + \frac{1}{2} M V^2 \tag{9.20}$$

Based on the definition of the position of the center of mass:

$$\vec{R} = \frac{1}{M} \sum_i m_i \vec{r}_i = \frac{1}{M} \sum_i m_i \left(\vec{R} + \vec{r}_i{}'\right) = \vec{R} + \frac{1}{M} \sum_i m_i \vec{r}_i{}' \tag{9.21}$$

we conclude that

$$\sum_i m_i \vec{r}_i{}' = 0 \tag{9.22}$$

The kinetic energy of the system is thus equal to

$$T = \frac{1}{2} M V^2 + \sum_i \frac{1}{2} m_i {v_i}'^2 \tag{9.23}$$

The change in the potential energy of the system when it moves from configuration 1 to configuration 2 is related to the work done by the forces acting on the system:

$$W_{12} = \sum_i \int_1^2 \vec{F}_{i,ext} \cdot d\vec{r}_i + \sum_i \int_1^2 \sum_{j \neq i} \vec{F}_{ij,\mathrm{int}} \cdot d\vec{r}_i \tag{9.24}$$

If we make the assumption that the forces, both internal and external, are derivable from potential functions, we can rewrite this expression as

$$W_{12} = -\sum_i \int_1^2 \vec{\nabla}(U_{i,ext}) \cdot d\vec{r}_i - \sum_i \int_1^2 \sum_{j \neq i} \vec{\nabla}\left(U_{ij,\text{int}}\right) \cdot d\vec{r}_i \tag{9.25}$$

The first term on the right-hand side can be evaluated easily:

$$\sum_i \int_1^2 \vec{\nabla}(U_{i,ext}) \cdot d\vec{r}_i = \sum_i U_{i,ext}\Big|_1^2 = \sum_i \{U_{i,ext}(2) - U_{i,ext}(1)\} \tag{9.26}$$

The second term can be rewritten as

$$\sum_i \int_1^2 \sum_{j \neq i} \vec{\nabla}\left(U_{ij,\text{int}}\right) \cdot d\vec{r}_i = \sum_i \int_1^2 \sum_{j<i} \left\{\vec{\nabla}\left(U_{ij,\text{int}}\right) \cdot d\vec{r}_i + \vec{\nabla}\left(U_{ji,\text{int}}\right) \cdot d\vec{r}_j\right\} = \sum_i \int_1^2 \sum_{j<i} \left\{\vec{\nabla}\left(U_{ij,\text{int}}\right) \cdot d\vec{r}_{ij}\right\} \tag{9.27}$$

Here we have used the fact that the internal force between i and j satisfy the following relation

$$\vec{F}_{ij,\text{int}} = -\vec{F}_{ji,\text{int}} \tag{9.28}$$

and thus

$$U_{ij,\text{int}} = -U_{ji,\text{int}} \tag{9.29}$$

The integral can be evaluated easily:

$$\sum_i \int_1^2 \sum_{j<i} \left\{\vec{\nabla}\left(U_{ij,\text{int}}\right) \cdot d\vec{r}_{ij}\right\} = \sum_i \sum_{j<i} \int_1^2 dU_{ij,\text{int}} = \sum_i \sum_{j<i} \left\{U_{ij,\text{int}}(2) - U_{ij,\text{int}}(1)\right\} \tag{9.30}$$

The work W_{12} is thus equal to

$$W_{12} = -\sum_i \{U_{i,ext}(2) - U_{i,ext}(1)\} - \sum_i \sum_{j<i} \left\{U_{ij,\text{int}}(2) - U_{ij,\text{int}}(1)\right\} \tag{9.31}$$

The total potential energy of the system U is defined as the sum of the internal and the external potential energy and is equal to

$$U = \sum_i U_{i,ext} + \sum_i \sum_{j<i} U_{ij,\text{int}} \tag{9.32}$$

The work done by all the forces to make the transition from configuration 1 to configuration 2 is

$$W_{12} = -U\big|_1^2 = -(U_2 - U_1) = U_1 - U_2 \tag{9.33}$$

-

Using the work-energy theorem we can conclude that

$$W_{12} = U_1 - U_2 = dT = T_2 - T_1 \tag{9.34}$$

or

$$E_1 = T_1 + U_1 = T_2 + U_2 = E_2 \tag{9.35}$$

We thus see that the total energy is conserved. If the system of particles is a rigid object, the components of the system will retain their relative positions, and the internal potential energy of the system will remain constant. We conclude that the total energy of the system has the following properties:

1. The total kinetic energy of the system is equal to the sum of the kinetic energy of a particle of mass M moving with the velocity of the center of mass and the kinetic energy of the motion of the individual particles relative to the center of mass.
2. The total energy for a conservative system is constant.

9.4.1 Example - Problem 9.21

A flexible rope of length 1.0 m slides from a frictionless table top as shown in Fig. 9.4 The rope is initially released from rest with 30 cm hanging over the edge of the table. Find the time at which the left end of the rope reaches the edge of the table.

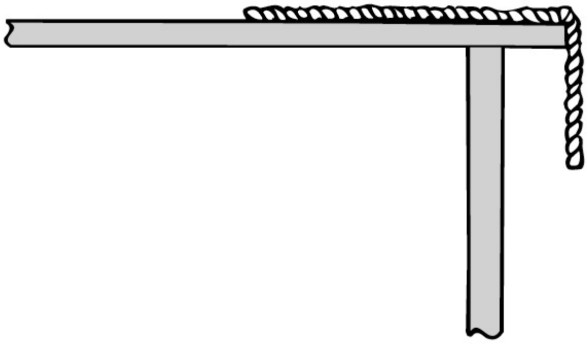


Figure 9.4: Problem 9.21.

Let us call x the length of rope hanging over the edge of the table, and ℓ the total length of the rope. The Lagrangian of the this system is equal to

$$L = \frac{1}{2}m\dot{x}^2 - (-\frac{1}{2}\frac{mg}{\ell}x^2) = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}\frac{mg}{\ell}x^2 \tag{9.36}$$

The Lagrange equation of motion can now be determined:

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = \frac{mg}{\ell}x - \frac{d}{dt}(m\dot{x}) = \frac{mg}{\ell}x - m\ddot{x} = 0 \tag{9.37}$$

The equation of motion is thus

$$\ddot{x} = \frac{g}{\ell}x \tag{9.38}$$

Let us look for solutions of the form

$$x = Ae^{\omega t} + Be^{-\omega t} \tag{9.39}$$

Putting this into equation of motion, we find

$$\omega = \sqrt{\frac{g}{\ell}} \tag{9.40}$$

The initial conditions are

$$x(t=0) = x_0 = 0.3\text{m} \tag{9.41}$$

and

$$v(t=0) = 0\text{m/s}. \tag{9.42}$$

From these we find that

$$A = B = \frac{1}{2}x_0 \tag{9.43}$$

We thus conclude that

$$x = x_0 \cosh(\omega t) \tag{9.44}$$

When $x = \ell$, the left end of the rope reaches the edge of the table, and the corresponding time is

$$t = \frac{1}{\omega}\cosh^{-1}\left(\frac{\ell}{x_0}\right) = 0.59\text{s}. \tag{9.45}$$

To verify our calculations, let us make sure that energy is conserved. Assume that the rope has a total mass M. We choose our coordinate system such that the vertical position coordinate on the surface of the table is 0; we also choose the surface of the table to be the plane in which the gravitational potential energy is equal to 0. To determine the change in the potential energy of the of the rope, we examine the change in the vertical position of its center of mass:

$$y_{i,cm} = \frac{1}{M}\left\{\left\{(L - x_0)\frac{M}{\ell}\right\}(0) + \left\{x_0\frac{M}{\ell}\right\}\left(-\frac{x}{2}\right)\right\} = -\frac{1}{2}\frac{{x_0}^2}{\ell} \tag{9.46}$$

and

$$y_{f,cm} = -\frac{\ell}{2} \tag{9.47}$$

The change in the potential energy is thus equal to

$$U = U_f - U_i = Mg\left(-\frac{1}{2}\ell\right) - Mg\left(-\frac{1}{2}\frac{{x_0}^2}{\ell}\right) = \frac{1}{2}MgL\left\{\frac{{x_0}^2}{\ell^2} - 1\right\} \tag{9.48}$$

Note: since the rope does not stretch, there is no change in the potential energy associated with the internal forces. To determine the change in the kinetic energy of the system, we need to determine the change in the velocity of the center of mass. The system is initially at rest, and the initial velocity of the center of mass is thus 0 (and so is its kinetic energy). The velocity of the system at the time the left end of the rope reaches the edge of the table can be found from the equations of motion:

$$v_{cm}(t) = \frac{dx(t)}{dt} = \frac{d}{dt}x_0\cosh\omega t = x_0\omega\sinh\omega t \tag{9.49}$$

When the rope reaches the edge of the table, the velocity of the center of mass is thus equal to

$$v_{cm}\left(t = \frac{1}{\omega}\cosh^{-1}\left(\frac{\ell}{x_0}\right)\right) = x_0\omega\sinh\left(\cosh^{-1}\left(\frac{\ell}{x_0}\right)\right) = x_0\omega\sinh\left(\pm\sinh^{-1}\left(\sqrt{\frac{\ell^2}{{x_0}^2} - 1}\right)\right) \tag{9.50}$$

This equation can be rewritten as

$$v_{cm}\left(t = \frac{1}{\omega}\cosh^{-1}\left(\frac{\ell}{x_0}\right)\right) = x_0\sqrt{\frac{g}{\ell}}\left(\sqrt{\frac{\ell^2}{{x_0}^2} - 1}\right) \tag{9.51}$$

The change in the kinetic energy of the system is thus equal to

$$\Delta T = T_f - T_i = \frac{1}{2}M{v_{cm,f}}^2 = \frac{1}{2}M{x_0}^2\frac{g}{\ell}\left(\frac{\ell^2}{{x_0}^2} - 1\right) = \frac{1}{2}\frac{M}{\ell}g\left(\ell^2 - {x_0}^2\right) = \frac{1}{2}Mg\ell\left(1 - \frac{{x_0}^2}{\ell^2}\right) = -\Delta U \tag{9.52}$$

9.5 Elastic and Inelastic Collisions

When two particles interact, the outcome of the interaction will be governed by the force law that describes the interaction. Consider an interaction force F_{int} that acts on a particle. The result of the interaction will be a change in the momentum of the particle since

$$F_{\text{int}} = ma = m\frac{dv}{dt} = \frac{dp}{dt} \tag{9.53}$$

If the interaction occurs over a short period of time, we expect to a change in the linear momentum of the particle:

$$\int_i^f \vec{F}_{\text{int}} dt = \int_i^f \left(\frac{d\vec{p}}{dt} \right) dt = \vec{p}_f - \vec{p}_i = \Delta \vec{p} \tag{9.54}$$

This relation shows us that if we know the force we can predict the change in the linear momentum, or if we measure the change in the linear momentum we can extract information about the force. We note that the change in the linear momentum provides us with information about the time integral of the force, not the force. Due to the importance of the time integral, it has received its own name, and is called the impulse P:

$$\vec{P} = \int_i^f \vec{F}_{\text{int}} dt \tag{9.55}$$

If we consider the effect of the interaction force on both particles we conclude that the change in the linear momentum is 0:

$$\Delta \vec{p} = \Delta \vec{p}_1 + \Delta \vec{p}_2 = \int_i^f \vec{F}_{12,\text{int}} dt + \int_i^f \vec{F}_{21,\text{int}} dt = \int_i^f \vec{F}_{12,\text{int}} dt - \int_i^f \vec{F}_{12,\text{int}} dt = 0 \tag{9.56}$$

This of course should be no surprise since when we consider both particles, the interaction force becomes an internal force and in the absence of external forces, linear momentum will be conserved.

Conservation of linear momentum is an important conservation law that restricts the possible outcomes of a collision. No matter what the nature of the collision is, if the initial linear momentum is non-zero, the final linear momentum will also be non-zero, and the system can not be brought to rest as a result of the collision. If the system is at rest after the collision, its linear momentum is zero, and the initial linear momentum must therefore also be equal to zero. Note that a zero linear momentum does not imply that all components of the system will be at rest; it only requires that the two object have linear momenta that are equal in magnitude but directed in opposite directions. The most convenient way to look at the collisions is in the center-of-mass frame. In the center-of-mass frame, the total linear momentum is equal to zero, and the objects will always travel in a co-linear fashion. This illustrated in Fig. 9.5.

We frequently divide collisions into two distinct groups:

1. **Elastic collisions**: collisions in which the total kinetic energy of the system is conserved. The kinetic energy of the objects will change as a result of the interaction, but the total kinetic energy will remain constant. The kinetic energy of one of the objects in general is a function of the masses of the two objects and the scattering angle.

2. **Inelastic collisions**: collisions in which the total kinetic energy of the system is not conserved. A totally inelastic collision is a collision in which the two objects stick together after the collision. The loss in kinetic energy is usually expressed in terms of the Q value, where $Q = K_f - K_i$:

    (a) $Q < 0$: **endoergic collision** (kinetic energy is lost)
    (b) $Q > 0$: **exoergic collision** (kinetic energy is gained)
    (c) $Q = 0$: **elastic collision** (kinetic energy is conserved).

In most inelastic collisions, a fraction of the initial kinetic energy is transformed into internal energy (for example in the form of heat, deformation, etc.). Another parameter that is frequently used to quantify the inelasticity of an inelastic collision is the coefficient of restitution e:

$$\varepsilon = \frac{|v_2 - v_1|}{|u_2 - u_1|} \tag{9.57}$$

where u are the velocities before the collision and v are the velocities after the collision. For a perfect elastic collision $e = 1$ and for a totally inelastic collision $e = 0$.

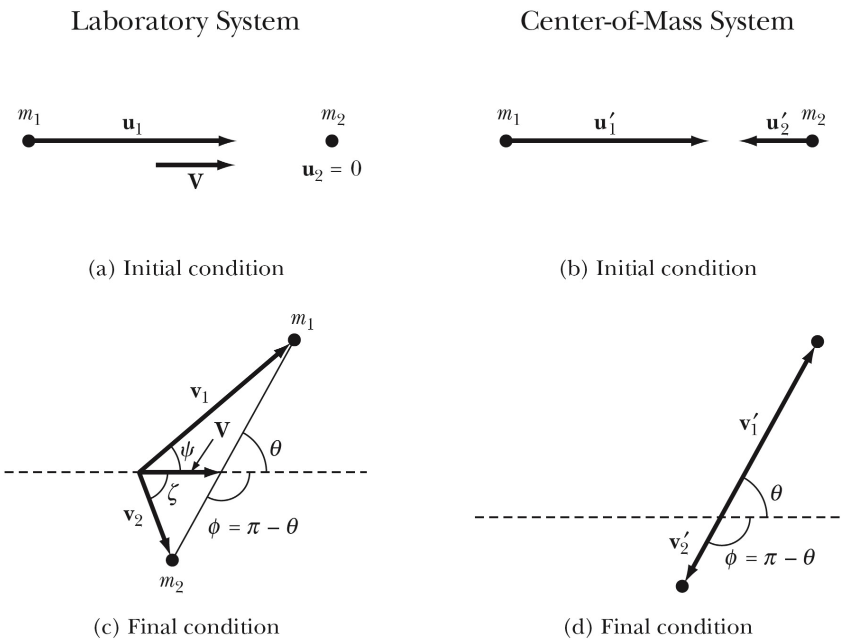


Figure 9.5: Two-dimensional collisions in the laboratory frame (left) and the center-of-mass frame (right).

One important issue we need to address when we focus on collisions is the issue of predictability. Let us consider what we know and what we need to know; we will assume that we are looking at a collision in the center-of-mass frame. Let us define the x axis to be the axis parallel to the direction of motion of the incident objects, and let us assume that the masses of the objects do not change. The unknown parameters are the velocities of the object; for the n-dimensional case, there will be $2n$ unknown. What do we know?

1. Conservation of linear momentum: this conservation law provides us with n equations with $2n$ unknown.
2. Conservation of energy: if the collision is elastic, this conservation law will provide us with 1 equation with $2n$ unknown.

For elastic collisions, we have $n + 1$ equations with $2n$ unknown. We immediately see that only for $n = 1$ the final state is uniquely defined. For inelastic collisions we have n equations with $2n$ unknown and we conclude that even for $n = 1$ the final state is undefined. When the final state is undefined, we need to know something about some of the final-state parameters to fix the others.

There are many applications of our collision theory. Consider one technique that can be used to study the composition of a target material. We use a beam of particles of known mass m_1 and kinetic energy $T_{initial}$ to bombard the target material and measure the energy of the elastically scattered projectiles at 90°. The measured kinetic energy depends on the mass of the target nucleus and is given by

$$T_{final} = \frac{{m_1}^2}{{m_1}^2 + {m_2}^2}\left[\left(\frac{m_2}{m_1}\right)^2 - 1\right]T_{initial} \tag{9.58}$$

By measuring the final kinetic energy, we can thus determine the target mass. Note: we need to make sure that the object we detect at 90° is the projectile. An example of an application of this is shown in Fig. 9.6.

9.6 Scattering Cross Section

We have learned a lot of properties of atoms and nuclei using elastic scattering of projectiles to probe the properties of the target elements. A schematic of the scattering process is shown in Fig. 9.7. In this Figure,

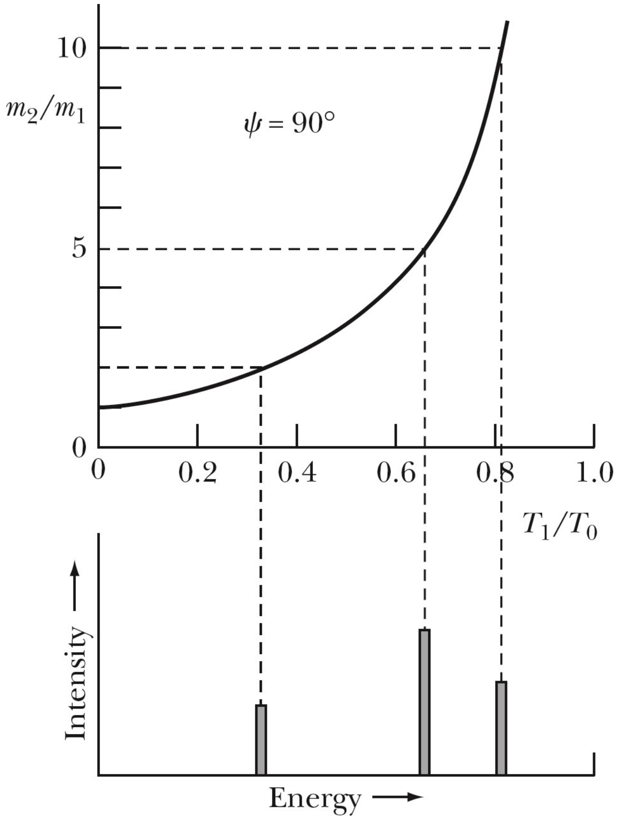


Figure 9.6: Energy spectrum of the scattered projectiles at 90°.

the incident particle is deflected (repelled) by the target particle. This situation will arise when we consider the scattering of positively charged nuclei. The parameter b is called the **impact parameter**. The impact parameter is related to the angular momentum of the projectile with respect to the target nucleus.

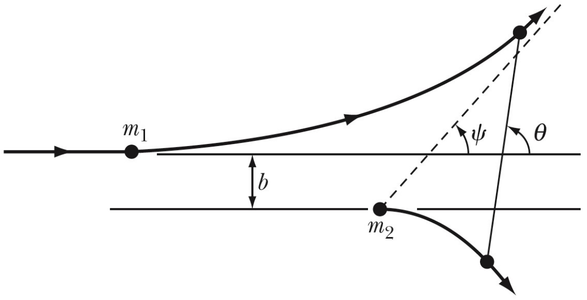


Figure 9.7: Scattering of projectile nuclei from target nuclei.

When we study the scattering process, we generally measure the intensity of the scattered particles as function of the scattering angle. The intensity distribution is expressed in terms of the differential cross section, which is defined as

$$\sigma(\theta) = \frac{\text{\#of interactions per target nucleus into an area } d\Omega' \text{ at } \theta}{\text{\# of incident particles per units area}} \tag{9.59}$$

There is a one-to-one correlation between the impact parameter b and the scattering angle θ (see Fig. 9.8). The one-to-one correlation is a direct consequence of the conservation of angular momentum. Assuming

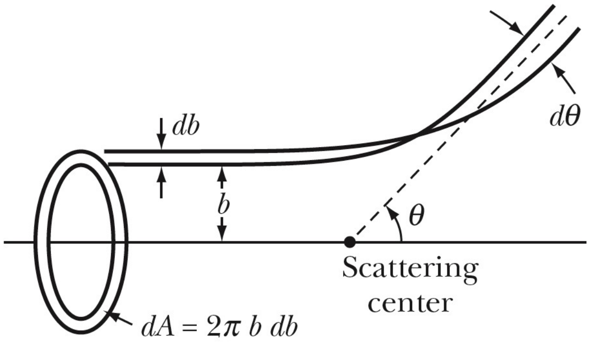


Figure 9.8: Correlation between impact parameter and scattering angle.

that the number of incident particles is conserved, the flux of incident particles with an impact parameter between b and $b + db$ is equal to

$$dI_{in} = I\{2\pi b db\} \tag{9.60}$$

must be the same as the number of particles scattered in the cone that is specified by the angle θ and width $d\theta$. The area of this cone is equal to

$$d\Omega' = 2\pi \sin\theta d\theta \tag{9.61}$$

The number of particles scattered into this cone will be

$$dI_{out} = -I\sigma(\theta) 2\pi \sin\theta d\theta \tag{9.62}$$

The minus sign is a result of the fact $db/d\theta < 0$. We thus conclude that

$$I\{2\pi b db\} = -I\sigma(\theta) 2\pi \sin\theta d\theta \tag{9.63}$$

or

$$\sigma(\theta) = \frac{\{2\pi b db\}}{\{-2\pi \sin\theta d\theta\}} = \frac{b}{\sin\theta}\left|\frac{db}{d\theta}\right| \tag{9.64}$$

The scattering angle θ is related to the impact parameter b and this relation can be obtained using the orbital motion we have discussed in this and in Chapter 8:

$$\theta = \int_{r_{\min}}^{\infty} \frac{\left(b/r^2\right) dr}{\sqrt{1 - (b^2/r^2) - (U/T_0{}')}} \tag{9.65}$$

The relation between the scattering angle θ and the impact parameter b depends on the potential U. For the important case of nuclear scattering, the potential varies as k/r. For this potential we can carry out the integration and determine the following correlation between the scattering angle and the impact parameter:

$$\cos\theta = \frac{(\kappa/b)}{\sqrt{1 + (\kappa/b)^2}} \tag{9.66}$$

where

$$\kappa = \frac{k}{2T_0{}'} \tag{9.67}$$

We can use this relation to calculate $db/d\theta$ and get the following differential cross section:

$$\sigma(\theta) = \left(\frac{k}{4T_0{}'}\right)^2 \frac{1}{\sin^4(\theta/2)} \tag{9.68}$$

We conclude that the intensity of scattered projectile nuclei will decrease when the scattering angle increases. If the energy of projectile nuclei is low enough, the measured angular distribution will agree with the so-called Rutherford distribution over the entire angular range, as was first shown by Geiger and Marsden in 1913 (see Fig. 9.9(left)).

Each trajectory of the projectile can be characterized by a distance of closest approach and there is a one-to-one correspondence between the scattering angle and this distance of closest approach. The smallest distance of closest approach occurs when the projectile is scattered backwards ($\theta = 180°$). The distance of closest approach decreases with increasing incident energy and the Rutherford formula indicates that the intensity should decrease as $1/T^2$. This was indeed observed, up to a maximum incident energy, beyond which the intensity dropped of much more rapidly than predicted by the Rutherford formula (see Fig. 9.9(right)). At this point, the nuclei approach each other so close that the strong attractive nuclear force starts to play a role, and the scattering is no longer elastic (the projectile nuclei may for example merge with the target nuclei).

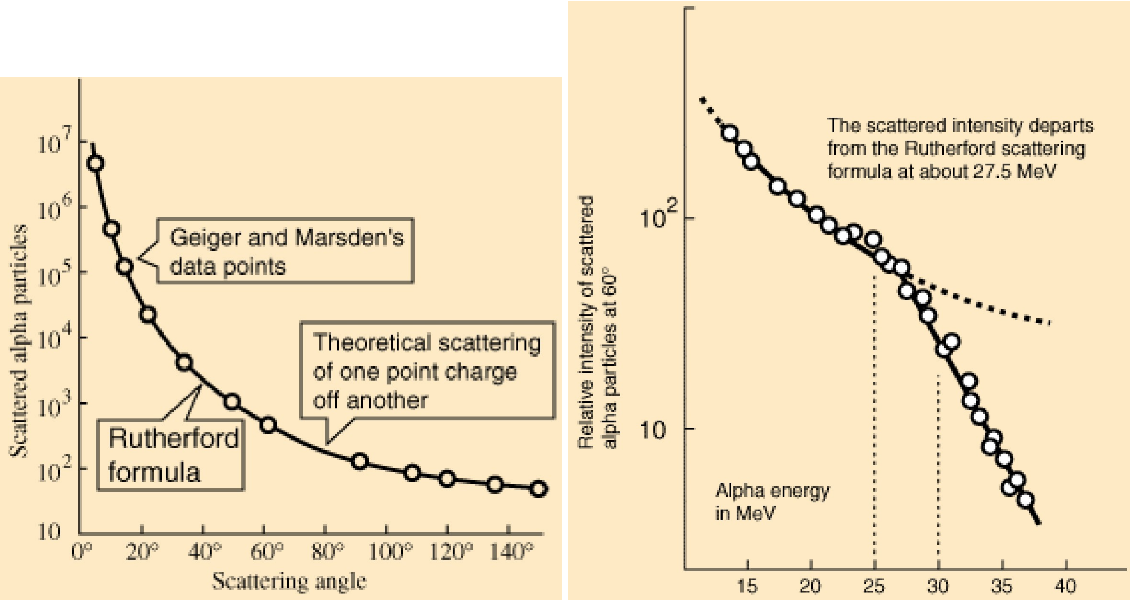


Figure 9.9: Measurement of the scattering of alpha particles from target nuclei. Figures taken from http://hyperphysics.phy-astr.gsu.edu/hbase/nuclear/rutsca2.html.

9.7 Rocket Motion

The motion of a rocket is a nice example of a system with a variable mass to which conservation of linear momentum can be applied. Suppose a rocket is flying through deep space (no friction force and no gravitational force). It is burning fuel. Suppose at some time t, the mass of the rocket is M. During a time interval dt, the mass of the rocket changes by dM:

$$M(t+dt) = M(t) + dM \tag{9.69}$$

Since the rocket is burning fuel, dM is negative, and the mass of the exhaust products is $-dM$. The result of the burning of fuel is a change in the velocity of the rocket:

$$v(t+dt) = v(t) + dv \tag{9.70}$$

If we consider our system to consist of the rocket and the exhaust generated during the time interval dt, we are dealing with a closed system. Since there are no external forces acting on the system, the total linear momentum of the system is conserved. The initial linear momentum of the system (at time t) is given by

$$p_i = M(t)v(t) \tag{9.71}$$

The final linear momentum of the system is given by

$$p_f = (M(t) + dM)(v(t) + dv) + (-dM)U \tag{9.72}$$

where U is the velocity of the exhaust. Conservation of linear momentum therefore requires that

$$M(t)v(t) = (M(t) + dM)(v(t) + dv) + (-dM)U \tag{9.73}$$

The exhaust velocity of the rocket depends on the design of the rocket engine. Suppose that for the engine used, the velocity of the exhaust relative to the engine is measured to be U_0. In the frame of reference in which the rocket is moving, the exhaust velocity is a function of both U_0 and the velocity of the rocket

$$U - U_0 = v(t) + dv \tag{9.74}$$

Using this expression, we can rewrite the expression for conservation of linear momentum as follows

$$M(t)v(t) = (M(t) + dM)(v(t) + dv) + (-dM)(v(t) + dv + U_0) \tag{9.75}$$

or

$$M(t)v(t) = M(t)(v(t) + dv) - dMU_0 \tag{9.76}$$

We conclude

$$dMU_0 = M(t)dv \tag{9.77}$$

Dividing each side of the previous equation by dt we obtain

$$\frac{dM}{dt}U_0 = M(t)\frac{dv}{dr} \tag{9.78}$$

Now consider the following substitutions:

1. $dM/dt = -R$ where R is the rate of fuel consumption.
2. $U_0 = -u$ where u is the (positive) velocity of the exhaust gases relative to the rocket.
3. dv/dt is the acceleration a of the rocket.

After making these substitutions, we obtain the "**first rocket equation**"

$$Ru = Ma \tag{9.79}$$

The mass used in the "first rocket equation" is of course time dependent (related to R). In order to find the velocity of the rocket after burning some fuel, we return to the differential equation previously discussed

$$-u\,dM = M\,dv \tag{9.80}$$

or

$$dv = -u\frac{dM}{M} \tag{9.81}$$

Integrating both sides gives

$$\int dv = v_f - v_i = -u\int\frac{dM}{M} = -u(\ln M_f - \ln M_i) = u\ln\frac{M_i}{M_f} \tag{9.82}$$

We conclude

$$v_f = v_i + u \ln \frac{M_i}{M_f} \tag{9.83}$$

which is the ”**second rocket equation”.**

10 Motion in a Non-Inertial Reference Frame

The laws of physics are only valid in inertial reference frames. However, it is not always easy to express the motion of interest in an inertial reference frame. Consider for example the motion of a book laying on top of a table. In a reference frame that is fixed with respect to the Earth, the motion is simple: if the book is at rest, it will remain at rest (here we assume that the surface of the table is horizontal). However, we do know that the Earth frame is not an inertial frame. In order to describe the motion of the book in an inertial frame, we need to take into account the rotation of the Earth around its axis, the rotation of the Earth around the Sun, the rotation of our solar system around the center of our galaxy, etc. etc. The motion of the book will all of a sudden be a lot more complicated!

For many experiments, the effect of the Earth not being an inertial reference frame is too small to be observed. Other effects, such as the tides, can only be explained if we take into consideration the non-inertial nature of the reference frame of the Earth and apply the laws of physics in an inertial frame.

10.0.1 Example: Problem 10.1

Calculate the centrifugal acceleration, due to the Earth's rotation, on a particle on the surface of the Earth at the equator. Compare this result with the gravitational acceleration. Computer also the centrifugal acceleration due to the motion of the Earth about the Sun and justify the remark made in the text that this acceleration may be neglected compared with the acceleration caused by axial rotation.

Consider the rotation of the Earth around its axis. The angular velocity associated with this rotation is equal to

$$\omega = 2\pi \text{ rad/day} = \frac{2\pi}{24 \times 3600} \text{ rad/s} \tag{10.1}$$

The radius of the Earth is 6.4×10^6 m. The acceleration of a mass located at the equator due to this rotation is equal to

$$a = \frac{v^2}{r} = \frac{\omega^2 r^2}{r} = r\omega^2 = 3.4 \times 10^{-2} \text{ m/s}^2 \tag{10.2}$$

Consider the rotation of the Earth around the Sun. The angular velocity associated with this rotation is equal to

$$\omega = 2\pi \text{ rad/year} = \frac{2\pi}{365.25 \times 24 \times 3600} \text{ rad/s} = 2.0 \times 10^{-7} \text{ rad/s} \tag{10.3}$$

The radius of the orbit of the Earth around the Sun is 1.5×10^{11} m. The acceleration of a mass located on the Earth due to the rotation of the Earth around the Sun is thus equal to

$$a = \frac{v^2}{r} = \frac{\omega^2 r^2}{r} = r\omega^2 = 6 \times 10^{-3} \text{ m/s}^2 \tag{10.4}$$

Finally, consider the rotation of the solar system around the center of the Milky Way. The velocity of the solar system is 230 km/s (2.3×10^5 m/s). The distance of the solar system to the center of the Milky Way is 8 kpc (8×10^{19} m). The acceleration of the solar system is thus equal to

$$a = \frac{v^2}{r} = 6.6 \times 10^{-10} \text{ m/s}^2. \tag{10.5}$$

Compared to the gravitational acceleration on the surface of the Earth (9.8 m/s^2), these centrifugal accelerations correspond to

$$\begin{array}{ll} 0.25\% & \text{for the rotation of the Earth around its axis} \\[1em] 0.061\% & \text{for the rotation of the Earth around the Sun} \\[1em] 6.7 \times 10^{-9}\% & \text{for the rotation of the solar system around the center of the Milky Way} \end{array} \tag{10.6}$$

10.1 Rotating Coordinate Systems

Consider the two coordinate systems shown in Fig. 10.1. The non-primed coordinates are the coordinates in the rotating frame, and the primed coordinates are the coordinates in the fixed coordinate system.

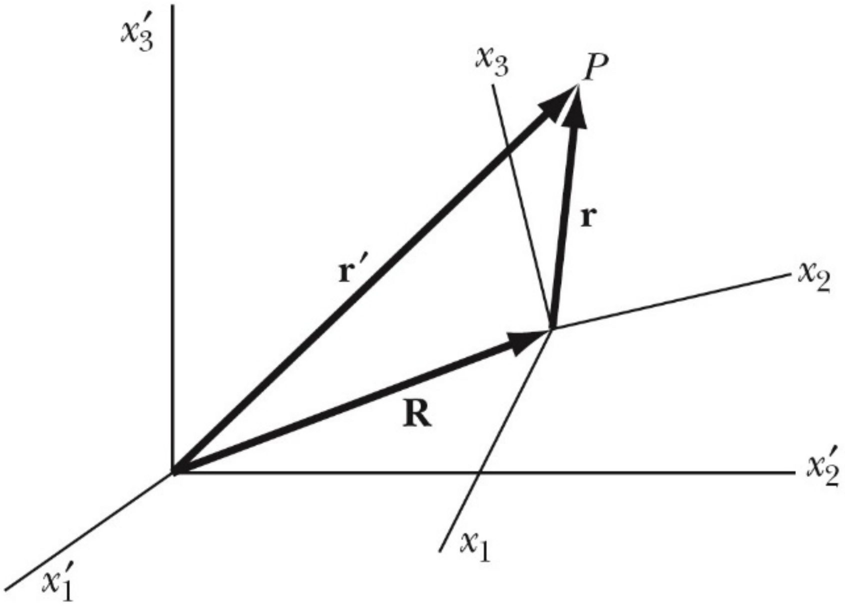


Figure 10.1: Fixed (primed) and rotating (non-primed) coordinate systems. The vector R specifies the origin of the rotating coordinate system in the fixed coordinate system.

Consider the motion of a point P. In the fixed coordinate system, the position of P is specified by the position vector $\vec{r}'$ and in the rotating coordinate system, its position is specified by the position vector $\vec{r}$. As can be seen in Fig. 10.1, these two vectors are related:

$$\vec{r}' = \vec{r} + \vec{R} \tag{10.7}$$

Consider what happens when the rotating coordinate system rotates by an infinitesimal angle $d\theta$. If point P is at rest in the rotating coordinate system, we will see the position of P in our fixed coordinate system change:

$$(d\vec{r})_{fixed} = d\vec{\theta} \times \vec{r} \tag{10.8}$$

Note: $(d\vec{r})_{fixed}$ indicates the the vector $d\vec{r}$ is measured in the fixed coordinate system.
If the rotation occurs during a period dt, we can rewrite the previous equation as

$$\left(\frac{d\vec{r}}{dt}\right)_{fixed} = \frac{d\vec{\theta}}{dt} \times \vec{r} = \vec{\omega} \times \vec{r} \tag{10.9}$$

To derive this relation, we have assumed that point P remains at rest in the rotating coordinate system. If point P is moving with respect to the rotating coordinate system, we need to add this contribution to the expression of the velocity of P in the fixed coordinate system:

$$\left(\frac{d\vec{r}'}{dt}\right)_{fixed} = \left(\frac{d\vec{r}}{dt}\right)_{rotating} + \vec{\omega} \times \vec{r}_{rotating} \tag{10.10}$$

This relation is valid for any vector, not just the position vector. If instead of the position vector we use the angular velocity vector, we find that

$$\left(\frac{d\vec{\omega}'}{dt}\right)_{fixed}=\left(\frac{d\vec{\omega}}{dt}\right)_{rotating}+\vec{\omega}\times\vec{\omega}=\left(\frac{d\vec{\omega}}{dt}\right)_{rotating} \tag{10.11}$$

This relation shows that the angular acceleration is the same in both reference frames.
In order to determine the velocity of point P in the fixed coordinate frame in terms of the velocity of point P in the rotating coordinate system, we have to go back to the correlation between the position vectors shown in Fig. 10.1. By differentiating the vectors with respect to time we obtain the following relation:

$$\left(\frac{d\vec{r}'}{dt}\right)_{fixed}=\left(\frac{d\vec{R}}{dt}\right)_{fixed}+\left(\frac{d\vec{r}}{dt}\right)_{fixed} \tag{10.12}$$

Using our expression for the velocity of P in the fixed coordinate system we find that

$$\left(\frac{d\vec{r}'}{dt}\right)_{fixed}=\left(\frac{d\vec{R}}{dt}\right)_{fixed}+\left(\frac{d\vec{r}}{dt}\right)_{rotating}+\vec{\omega}\times\vec{r} \tag{10.13}$$

This equation can also be rewritten as

$$\vec{v}_f=\left(\frac{d\vec{r}'}{dt}\right)_{fixed}=\left(\frac{d\vec{R}}{dt}\right)_{fixed}+\left(\frac{d\vec{r}}{dt}\right)_{rotating}+\vec{\omega}\times\vec{r}=\vec{V}+\vec{v}_r+\vec{\omega}\times\vec{r} \tag{10.14}$$

where

$$\vec{v}_f=\left(\frac{d\vec{r}'}{dt}\right)_{fixed}=\text{velocity of P in the fixed frame} \tag{10.15}$$

$$\vec{V}=\left(\frac{d\vec{R}}{dt}\right)_{fixed}=\text{ velocity of the origin of the rotating frame} \tag{10.16}$$

$$\vec{v}_r=\left(\frac{d\vec{r}}{dt}\right)_{rotating}=\text{velocity of P in the rotating frame} \tag{10.17}$$

$$\vec{\omega}\times\vec{r}=\text{ velocity of P due to the rotation of the axes} \tag{10.18}$$

10.2 "Newton's Law" in Rotating Reference Frames

Consider the situation in which an external force $\vec{F}$ is acting on P. Only in the fixed reference frame can we use Newton's second law to determine the corresponding acceleration of P:

$$\vec{a}_f=\left(\frac{d\vec{v}_f}{dt}\right)_{fixed}=\frac{\vec{F}}{m} \tag{10.19}$$

Another expression for the acceleration of P can be obtained by differentiating the velocity-relation obtained in the previous section with respect to time:

$$\begin{aligned}\vec{a}_f &=\left(\tfrac{d\vec{v}_f}{dt}\right)_{fixed}=\left(\tfrac{d\vec{V}}{dt}\right)_{fixed}+\left(\tfrac{d\vec{v}_r}{dt}\right)_{fixed}+\left(\tfrac{d\vec{\omega}}{dt}\right)_{fixed}\times\vec{r}+\vec{\omega}\times\left(\tfrac{d\vec{r}}{dt}\right)_{fixed}=\\ &=\left(\tfrac{d\vec{V}}{dt}\right)_{fixed}+\left\{\left(\tfrac{d\vec{v}_r}{dt}\right)_{rotating}+\vec{\omega}\times\vec{v}_r\right\}+\left(\tfrac{d\vec{\omega}}{dt}\right)_{fixed}\times\vec{r}+\vec{\omega}\times\left\{\left(\tfrac{d\vec{r}}{dt}\right)_{rotating}+\vec{\omega}\times\vec{r}\right\}=\\ &=\left(\tfrac{d\vec{V}}{dt}\right)_{fixed}+\left(\tfrac{d\vec{v}_r}{dt}\right)_{rotating}+2\vec{\omega}\times\vec{v}_r+\dot{\vec{\omega}}\times\vec{r}+\vec{\omega}\times\{\vec{\omega}\times\vec{r}\}\end{aligned} \tag{10.20}$$

An observer in the rotating reference frame will observe an acceleration

$$\vec{a}_r = \left(\frac{d\vec{v}}{dt}\right)_{rotating} \tag{10.21}$$

This acceleration is certainly not equal to $\vec{F}/m$, but the previous relations can be used to express the acceleration in the rotating reference frame in terms of the acceleration in the fixed reference frame:

$$\vec{a}_r = \vec{a}_f - \left(\frac{d\vec{V}}{dt}\right)_{fixed} - 2\vec{\omega}\times\vec{v}_r - \dot{\vec{\omega}}\times\vec{r} - \vec{\omega}\times\{\vec{\omega}\times\vec{r}\} \tag{10.22}$$

This relation immediately shows what has been repeated already many times: the acceleration of an object at P will be the same in two reference frames, only if one frame does not rotate with respect to the other frame (that is $\omega = 0$ rad/s and $d\omega/dt = 0$ rad/s^2) and if the two reference frames do not accelerate with respect to each other.
In order to explore the implication of the relation between the acceleration of P in the rotating and in the fixed coordinate frames, we assume for the moment that the origin of the rotating reference frame is not accelerating with respect to the origin of the fixed reference frame ($dV/dt = 0$ m/s^2), and that the axes of the rotating reference frame are rotating with a constant angular velocity ($d\omega/dt = 0$ rad/s^2). Under these assumptions, we find that the acceleration of P in the rotating reference frame is equal to

$$\vec{a}_r = \vec{a}_f - 2\vec{\omega}\times\vec{v}_r - \vec{\omega}\times\{\vec{\omega}\times\vec{r}\} \tag{10.23}$$

The second and third terms on the right-hand side are non-inertial terms that are introduced to correct the real force $\vec{F}$ in order to be able to use Newton-like laws in the rotating frame:

$$\vec{F}_{eff} = m\vec{a}_f - 2m\vec{\omega}\times\vec{v}_r - m\vec{\omega}\times\{\vec{\omega}\times\vec{r}\} \tag{10.24}$$

Using this effective force, an observer in the rotating frame will be able to determine the acceleration in the rotating frame by dividing this effective force by the mass of the object.
The second term on the right-hand side of Eq. 10.24 is called the **Coriolis force**, and the last term on the right-hand side is called the **centripetal force**. Both of these forces are however not real forces; they are introduced in order to be able to use an equation similar to Newton's second law in non-inertial reference frames. When we try to describe an object on the surface of the Earth, we need to take the effects of these artificial forces into consideration. In the next two sections we will focus on these two forces in some detail.

10.2.1 Example: Problem 10.6

A bucket of water is set spinning about its symmetry axis. Determine the shape of the water in the bucket.

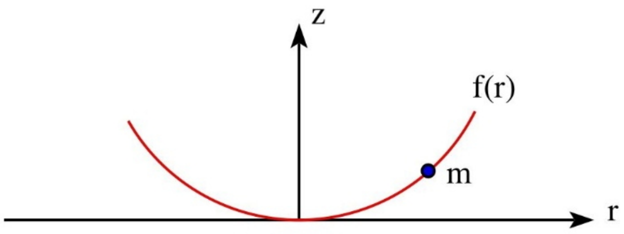


Figure 10.2: A spinning bucket of water.

Consider a small mass m on the surface of the water, shown in Fig. 10.2. From Eq. (10.25) in our text book we get

$$\vec{F}_{eff} = \vec{F} - m\ddot{\vec{R}}_f - m\dot{\vec{\omega}}\times\vec{r} - m\vec{\omega}\times(\vec{\omega}\times\vec{r}) - 2m\vec{\omega}\times\vec{v}_r \tag{10.25}$$

In the rotating frame, the mass is at rest; thus

$$\vec{F}_{eff} = 0 \tag{10.26}$$

The force $\vec{F}$ will consist of gravity and the force due to the pressure gradient, which is normal to the surface in equilibrium. Since

$$\ddot{\vec{R}}_f = \dot{\vec{\omega}} = \vec{v}_r = 0 \tag{10.27}$$

we now have

$$0 = m\vec{g} + \vec{F}_p - m\vec{\omega} \times (\vec{\omega} \times \vec{r}) \tag{10.28}$$

where $\vec{F} + P$ is due to the pressure gradient.

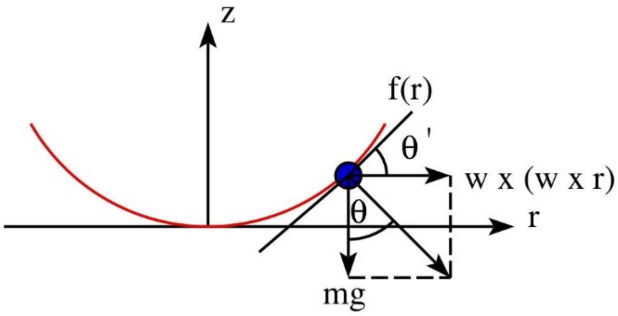


Figure 10.3: A spinning bucket of water.

Since $F_{eff} = 0$, the sum of the gravitational and centrifugal forces must also be normal to the surface. Thus $\theta' = \theta$.

$$\tan\theta' = \tan\theta = \frac{\omega^2 r}{g} \tag{10.29}$$

but

$$\tan\theta' = \frac{dz}{dr} \tag{10.30}$$

Thus

$$z = \frac{\omega^2}{2g} r^2 + \text{constant} \tag{10.31}$$

10.3 The Centripetal Force

The surface of the Earth is a non-inertial reference frame. The biggest deviation from good "inertial" behavior is due to the rotation of the Earth around its axis. In the current discussion we will thus ignore the motion of the Earth around the Sun, the motion of the solar system in our galaxy, etc. etc.

Consider a pendulum at rest in our rotating reference frame, which is at rest with respect to the surface of the Earth. Since the pendulum is at rest in this rotating reference frame, its velocity v_r in this frame is zero. The effective force seen by the pendulum is thus equal to

$$\vec{F}_{eff} = m\vec{a}_f - m\vec{\omega} \times \{\vec{\omega} \times \vec{r}\} = m\{\vec{g}_0 - \vec{\omega} \times \{\vec{\omega} \times \vec{r}\}\} \tag{10.32}$$

The direction of $\vec{g}_0$ in this equation is directed towards the center of the Earth, while the direction of the non-inertial correction term is radially outwards (see Fig. 10.4). If the angle between the position vector $\vec{r}$ and the rotation axis is equal to θ, we find the magnitude of the correction term is equal to

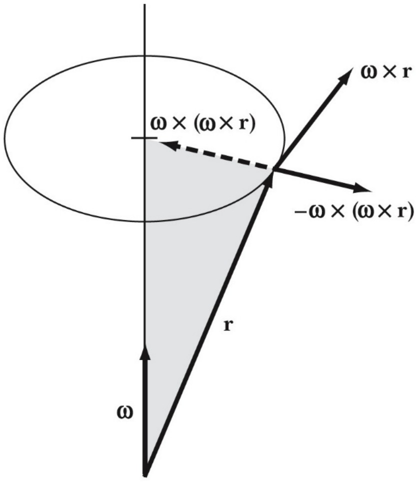


Figure 10.4: Direction of the centripetal correction term.

$$\left|\vec{\omega}\times\{\vec{\omega}\times\vec{r}\}\right| = \omega^2 r\sin\theta \tag{10.33}$$

The effect of this correction is that the equilibrium position of the pendulum (the position in which the arm of the pendulum is parallel to the direction of the net force) is changed, and the arm of the pendulum no longer points towards the center of the Earth (see Fig. 10.5). The direction of the gravitational acceleration, as measured by an observer in the rotating reference frame, is thus equal to

$$\vec{g} = \vec{g}_0 - \vec{\omega}\times\{\vec{\omega}\times\vec{r}\} \tag{10.34}$$

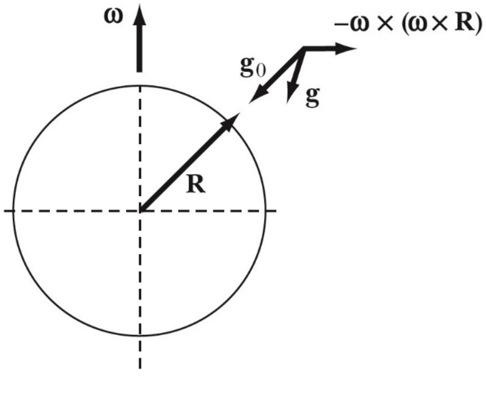


Figure 10.5: Effect of the centripetal term on a pendulum located on the surface of the Earth.

The centripetal correction changes both the magnitude of the observed acceleration and its direction. The angle between the direction of g_0 and the direction of g can be found easily (see Fig. 10.6):

$$\Delta\theta = \theta - \alpha = \theta - \mathrm{atan}\left(\frac{g_0\sin\theta - \omega^2 R\sin\theta}{g_0\cos\theta}\right) = \theta - \mathrm{atan}\left(\left(1-\frac{\omega^2}{g_0}\right)\tan\theta\right) \tag{10.35}$$

The same result could have been obtained if we had solved this problem in a non-rotating frame. Consider a simple pendulum of mass m attached a string. There are two forces acting on this mass: the tension $\vec{T}$ in the string and the gravitational force $\vec{F}_g$. An observer in the inertial frame will observe that mass m carries

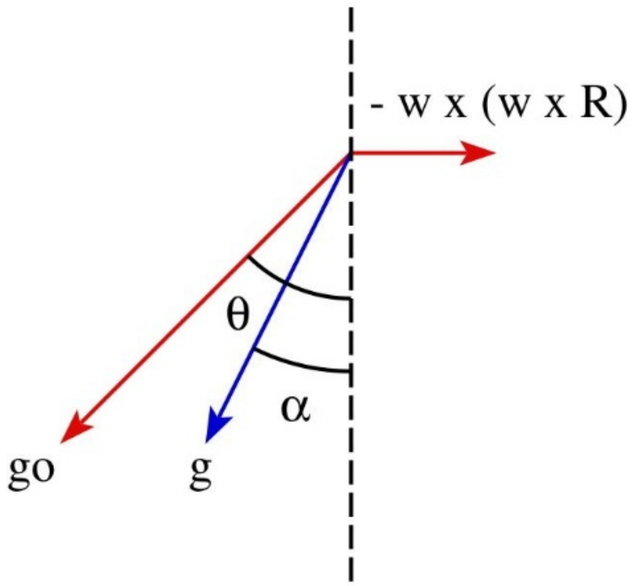


Figure 10.6: Direction of net gravitational acceleration.

out circular motion, with a radius $R\ \sin\theta$, and knows that there must be a net force acting on it, pointing towards the rotation axis. This force must have a magnitude of

$$F_r = m\frac{v^2}{R\sin\theta} = m\frac{\left(\frac{2\pi R\sin\theta}{T}\right)^2}{R\sin\theta} = m\omega^2 R\sin\theta \tag{10.36}$$

This force must be generated by the component of the tension and the gravitational force in this direction. We must thus require that (see Fig. 10.7):

$$mg_0\sin\theta - T\sin\alpha = m\omega^2 R\sin\theta \tag{10.37}$$

or

$$T\sin\alpha = mg_0\sin\theta - m\omega^2 R\sin\theta \tag{10.38}$$

The net force in the direction perpendicular to the plan of rotation must be zero, and we must thus require that

$$T\cos\alpha = mg\cos\theta \tag{10.39}$$

Combining these two equations, we obtain the following relation between the angles:

$$\tan\alpha = \frac{T\sin\alpha}{T\cos\alpha} = \frac{mg_0\sin\theta - m\omega^2 R\sin\theta}{mg_0\cos\theta} = \tan\left(\left(1-\frac{\omega^2}{g_0}\right)\tan\theta\right) \tag{10.40}$$

which is the same result we obtained previously.

10.3.1 Example: Problem 10.20

Calculate the effective gravitational field vector $\vec{g}$ at the Earth's surface at the poles and at the equator. Take the difference in the equatorial (6378 km) and polar (6357 km) radius into account.

The mass of the Earth is 5.976×10^{24} kg. At the pole, $\vec{v}_r = 0$ m/s and the Coriolis force is equal to 0 N. The magnitude of the gravitational field is thus equal to

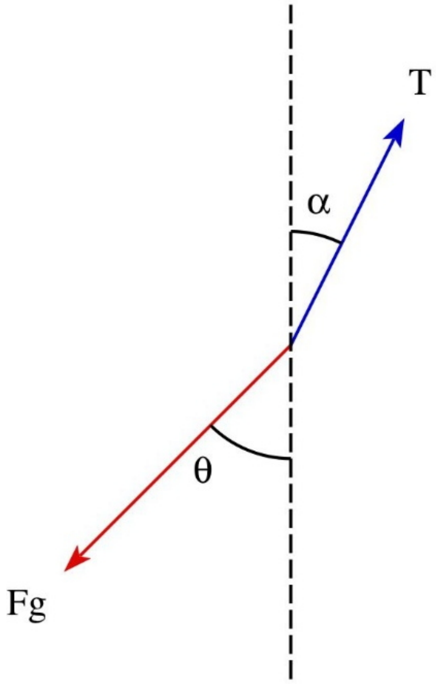


Figure 10.7: Pendulum in inertial frame.

$$g_{pole} = G\frac{M}{R_{pole}^2} = 9.866 \text{ m/s}^2 \tag{10.41}$$

On the equator, the Coriolis force is not equal to 0 N, and the magnitude of the gravitational field is equal to

$$g_{pole} = G\frac{M}{R_{eq}^2} - \omega^2 R_{eq} = 9.768 \text{ m/s}^2 \tag{10.42}$$

10.4 Coriolis Force

The Coriolis force is responsible for the deflection of objects moving in a rotating coordinate system. The force is proportional to the vector product of the angular velocity vector of the rotating coordinate system (as measured by an observer in a fixed coordinate system) and the velocity vector of the object in the rotating coordinate frame:

$$\vec{F}_{Coriolis} = -2m(\vec{\omega} \times \vec{v}_r) \tag{10.43}$$

The effect of the Coriolis force on the motion of an object is illustrated in Fig. 10.8. Note that the deflection depends on the z component of the angular velocity vector, which is perpendicular to the surface of the Earth. The z component reaches a maximum value at the North pole, and is zero at the equator.

As a result of the Coriolis force, air flowing from West to East towards a region of low pressure will be deflected to the South on the Northern hemisphere. Air approaching the low from the East will be deflected to the North. On the Northern hemisphere we expect that the air is flowing counter clockwise around an area of low pressure; in the same manner we can show that air flows clockwise around an area of high pressure. A lot about the weather can be understood on the basis of these observations. See for example the forecast map shown in Fig. 10.9. The position of the high-pressure system over Michigan will bring cold air from Canada to Rochester (since the circulation around the high is in the clockwise direction). We thus expect the winds to be from the North. Once the high passes Rochester, the wind should come from

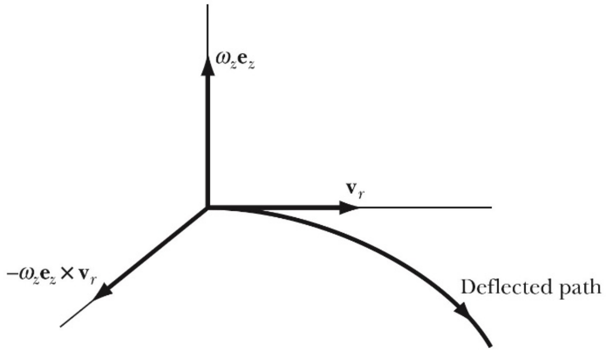


Figure 10.8: Deflection of a moving object as a result of the Coriolis force.

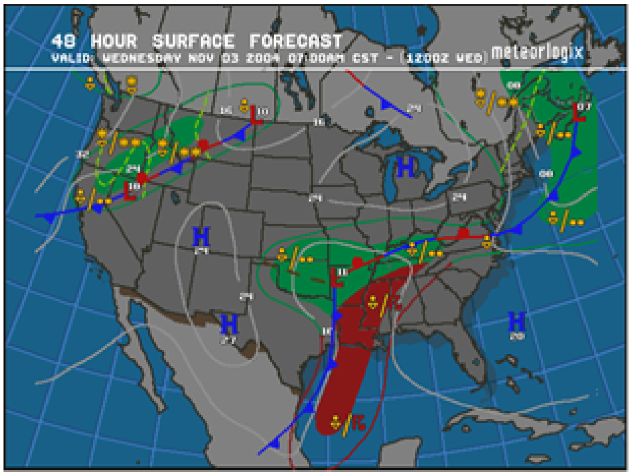


Figure 10.9: Forty-eight hour forecast map for Wednesday November 3, 2004, at 1200 Z (http://www.aopa.org/members/wx/focpage.cfm?sfcmap=0700d485).

the South, bringing us higher temperatures. The low in the South of the USA will pull in moisture from the gulf of Mexico and rain and thunder can be expected in the region in front of the low (since this is the region where moisture of the golf of Mexico will go as a result of the counter-clockwise flow around the low).

10.4.1 Example: Problem 10.18

A British warship fires a projectile due south near the Falkland Islands during World War I at a latitude 50° South. If the shells are fired at a 37° elevation with a speed of $v_0 = 800$ m/s, by how much do the shells miss their target and in what direction?

Consider the rotating reference frame. Its origin is located at the location of the warship and the definition of its axes is shown in Fig. 10.10.
The angular velocity at the location of the warship is

$$\vec{\omega} = \begin{pmatrix} -\omega \cos \alpha \\ -\omega \sin \alpha \\ 0 \end{pmatrix} \tag{10.44}$$

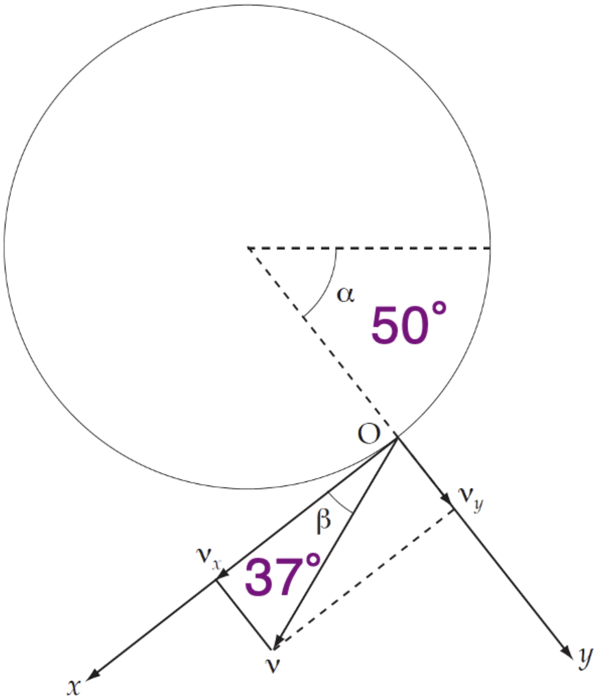


Figure 10.10: The coordinate system used to describe the motion of the projectile.

where $\alpha = 50°$. The velocity of the projectile in the rotating reference frame is

$$\vec{v} = \begin{pmatrix} v_0 \cos\beta \\ v_0 \sin\beta - gt \\ 0 \end{pmatrix} \tag{10.45}$$

where $\beta = 37°$. The acceleration due to the Coriolis force is equal to

$$a_c = 2\vec{v} \times \vec{\omega} = (-2v_0\omega\cos\beta\sin\alpha + 2(v_0\sin\beta - gt)\omega\cos\alpha)\hat{z} \tag{10.46}$$

The Coriolis force only influences the motion along the z axis. The projectile has an initial velocity of 0 m/s along the z axis. The z component of its velocity at a later time t is equal to

$$v_c = \int_0^t a_c dt = -2v_0\omega t(\cos\beta\sin\alpha - \sin\beta\cos\alpha) - gt^2\omega\cos\alpha \tag{10.47}$$

The z deflection of the projectile due to the Coriolis force is equal to

$$z_c = \int_0^t v_c dt = -v_0\omega t^2 \sin(\alpha - \beta) - \frac{1}{3}gt^3\omega\cos\alpha \tag{10.48}$$

The time at which the projectile hits the surface can be found by looking at its motion on the y axis since the motion along the y axis is not effected by the Coriolis force. This time is equal to

$$t = 2\frac{v_0\sin\beta}{g} \tag{10.49}$$

Note that the projectile reaches its highest position along the y axis at time $t = v_o \sin\beta/g$.
Using the time obtained from Eq. 10.49, the deflection of the projectile in the z direction can be determined: $z_c = 260$ m.

11 Dynamics of Rigid Bodies

A rigid body is a collection of particles with fixed relative positions, independent of the motion carried out by the body. The dynamics of a rigid body has been discussed in our introductory courses, and the techniques discussed in these courses allow us to solve many problems in which the motion can be reduced to two-dimensional motion. In this special case, we found that the angular momentum associated with the rotation of the rigid object is directed in the same direction as the angular velocity:

$$\vec{L} = I\vec{\omega} \tag{11.1}$$

In this equation, I is the moment of inertia of the rigid body which was defined as

$$I = \sum_i m_i {r_i}^2 \tag{11.2}$$

where r_i is the distance of mass m_i from the rotation axis. We also found that the kinetic energy of the body, associated with its rotation, is equal to

$$T = \frac{1}{2} I \omega^2 \tag{11.3}$$

The complexity of the motion increases when we need three dimensions to describe the motion. There are many different ways to describe motion in three dimensions. One common method is to describe the motion of the center of mass (in a fixed coordinate system) and to describe the motion of the components around the center of mass (in the rotating coordinate system).

11.1 The Inertia Tensor

In Chapter 10 we derived the following relation between the velocity of a particle in the fixed reference frame, v_f, and its velocity in the rotating reference frame v_r:

$$\vec{v}_f = \vec{V} + \vec{v}_r + \vec{\omega} \times \vec{r} \tag{11.4}$$

If we assume that the rotating frame is fixed to the rigid body, then $v_r = 0$.
The total kinetic energy of the rigid body is the sum of the kinetic energies of each component of the rigid body. Thus

$$\begin{aligned} T &= \textstyle\sum_\alpha \left\{ \frac{1}{2} m_\alpha {v_\alpha}^2 \right\} = \frac{1}{2} \sum_\alpha m_\alpha \left(\left\{ \vec{V} + \vec{\omega} \times \vec{r}_\alpha \right\} \cdot \left\{ \vec{V} + \vec{\omega} \times \vec{r}_\alpha \right\} \right) = \\ &= \textstyle\frac{1}{2} \sum_\alpha m_\alpha \left(V^2 + 2\vec{V} \cdot \{ \vec{\omega} \times \vec{r}_\alpha \} + \{ \vec{\omega} \times \vec{r}_\alpha \} \cdot \{ \vec{\omega} \times \vec{r}_\alpha \} \right) \end{aligned} \tag{11.5}$$

Let us now examine the three terms in this expression:

$$\frac{1}{2} \sum_\alpha m_\alpha V^2 = \frac{1}{2} V^2 \sum_\alpha m_\alpha = \frac{1}{2} M V^2 \tag{11.6}$$

$$\frac{1}{2} \sum_\alpha m_\alpha \left(2\vec{V} \cdot \{ \vec{\omega} \times \vec{r}_\alpha \} \right) = \vec{V} \cdot \left\{ \vec{\omega} \times \sum_\alpha (m_\alpha \vec{r}_\alpha) \right\} = \vec{V} \cdot \left\{ \vec{\omega} \times \left(M\vec{R} \right) \right\} = 0 \tag{11.7}$$

$$\begin{aligned} \textstyle\frac{1}{2} \sum_\alpha m_\alpha \left(\{ \vec{\omega} \times \vec{r}_\alpha \} \cdot \{ \vec{\omega} \times \vec{r}_\alpha \} \right) &= \textstyle\frac{1}{2} \sum_\alpha m_\alpha \left(\omega^2 {r_\alpha}^2 - \{ \vec{\omega} \cdot \vec{r}_\alpha \}^2 \right) = \\ &= \textstyle\frac{1}{2} \sum_\alpha m_\alpha \left(\left[\sum_i {\omega_i}^2 \right] \left[\sum_k {r_{\alpha,k}}^2 \right] - \left[\sum_i (\omega_i r_{\alpha,i}) \right] \left[\sum_j \left(\omega_j r_{\alpha,j} \right) \right] \right) = \\ &= \textstyle\frac{1}{2} \sum_{i,j} \left\{ \sum_\alpha m_\alpha \left(\delta_{ij} \left[\sum_k {r_{\alpha,k}}^2 \right] - r_{\alpha,i} r_{\alpha,j} \right) \right\} \omega_i \omega_j = \frac{1}{2} \sum_{i,j} I_{ij} \omega_i \omega_j \end{aligned} \tag{11.8}$$

The second term is zero, if we choose the origin of the rotating coordinate system to coincide with the center of mass of the rigid object.

Using the previous expressions, we can now rewrite the total kinetic energy of the rigid object as

$$T = \frac{1}{2}MV^2 + \frac{1}{2}\sum_{i,j} I_{ij}\omega_i\omega_j = T_{CM} + T_{rot} \tag{11.9}$$

The quantity I_{ij} is called the **inertia tensor**, and is a 3 x 3 matrix:

$$\{I\} = \begin{pmatrix} \sum_\alpha m_\alpha \left(r_{\alpha,2}{}^2 + r_{\alpha,3}{}^2\right) & -\sum_\alpha m_\alpha r_{\alpha,1} r_{\alpha,2} & -\sum_\alpha m_\alpha r_{\alpha,1} r_{\alpha,3} \\ -\sum_\alpha m_\alpha r_{\alpha,2} r_{\alpha,1} & \sum_\alpha m_\alpha \left(r_{\alpha,1}{}^2 + r_{\alpha,3}{}^2\right) & -\sum_\alpha m_\alpha r_{\alpha,2} r_{\alpha,3} \\ -\sum_\alpha m_\alpha r_{\alpha,3} r_{\alpha,1} & -\sum_\alpha m_\alpha r_{\alpha,3} r_{\alpha,2} & \sum_\alpha m_\alpha \left(r_{\alpha,1}{}^2 + r_{\alpha,2}{}^2\right) \end{pmatrix} \tag{11.10}$$

Based on the definition of the inertia tensor, we make the following observations:

1. The tensor is symmetric: $I_{ij} = I_{ji}$. Of the 9 parameters, only 6 are free parameters.
2. The non-diagonal tensor elements are called **products of inertia.**
3. The diagonal tensor elements are the moments of inertia with respect to the three coordinate axes of the rotating frame.

11.2 Angular Momentum

The total angular momentum L of the rotating rigid object is equal to the vector sum of the angular momenta of each component of the rigid object. The i^{th} component of L is equal to

$$\begin{aligned} L_i &= \sum_\alpha (\vec{r}_\alpha \times \vec{p}_\alpha)_i = \sum_\alpha (\vec{r}_\alpha \times m_\alpha (\vec{\omega} \times \vec{r}_\alpha))_i = \sum_\alpha m_\alpha (\vec{r}_\alpha \times \vec{\omega} \times \vec{r}_\alpha)_i = \sum_\alpha m_\alpha \left(r_\alpha{}^2 \vec{\omega} - \vec{r}_\alpha (\vec{r}_\alpha \cdot \vec{\omega})\right)_i = \\ &= \sum_\alpha m_\alpha \left\{r_\alpha{}^2 \omega_i - r_{\alpha,i} \sum_j \left(r_{\alpha,j}\omega_j\right)\right\} = \sum_j \omega_j \sum_\alpha m_\alpha \left\{r_\alpha{}^2 \delta_{ij} - r_{\alpha,i} r_{\alpha,j}\right\} = \sum_j I_{ij}\omega_j \end{aligned} \tag{11.11}$$

This equation clearly shows that the angular momentum is in general not parallel to the angular velocity. An example of a system where the angular momentum is directed in a direction different from the direction of the angular velocity is shown in Fig. 11.1.
The rotational kinetic energy can also be rewritten in terms of the angular momentum:

$$T_{rot} = \frac{1}{2}\sum_{i,j} I_{ij}\omega_i\omega_j = \frac{1}{2}\sum_i \omega_i \left(\sum_j I_{ij}\omega_j\right) = \frac{1}{2}\sum_i \omega_i L_i = \frac{1}{2}\left(\vec{\omega} \cdot \vec{L}\right) \tag{11.12}$$

11.3 Principal Axes

We always have the freedom to choose our coordinate axes such that the problem we are trying solve is simplified. When we are working on problems that involve the use of the inertia tensor, we can obtain a significant simplification if we can choose our coordinate axes such that the non-diagonal elements are 0. In this case, the inertia tensor would be equal to

$$\{I\} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \tag{11.13}$$

For this inertia tensor we get the following relation between the angular momentum and the angular velocity:

$$L_i = I_i\omega_i \tag{11.14}$$

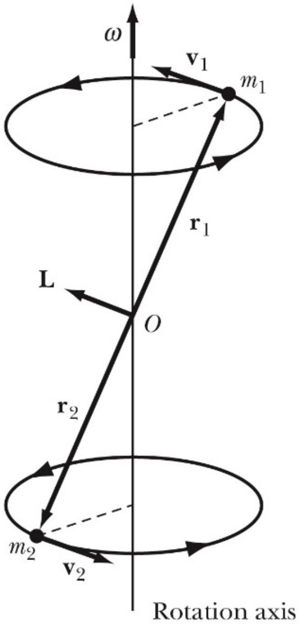


Figure 11.1: A rotating dumbbell is an example of a system in which the angular velocity is not parallel to the angular momentum.

The rotational kinetic energy is equal to

$$T_{rot} = \frac{1}{2}\sum_i I_i {\omega_i}^2 \tag{11.15}$$

The axes for which the non-diagonal matrix elements vanish are called the **principal axes of inertia**.
The biggest problem we are facing is how do we determine the proper coordinate axes? If the angular velocity vector is directed along one of the three coordinate axes that would get rid of the non-diagonal inertia tensor elements, and we expect to see the following relation between the angular velocity vector and the angular momentum:

$$\vec{L} = I\vec{\omega} \tag{11.16}$$

Substituting the general form of the inertia tensor into this expression, we must require that

$$\begin{aligned} L_1 &= I\omega_1 = I_{11}\omega_1 + I_{12}\omega_2 + I_{13}\omega_3 \\ L_2 &= I\omega_2 = I_{21}\omega_1 + I_{22}\omega_2 + I_{23}\omega_3 \\ L_3 &= I\omega_3 = I_{31}\omega_1 + I_{32}\omega_2 + I_{33}\omega_3 \end{aligned} \tag{11.17}$$

This set of equations can be rewritten as

$$\begin{aligned} (I_{11} - I)\,\omega_1 + I_{12}\omega_2 + I_{13}\omega_3 &= 0 \\ I_{21}\omega_1 + (I_{22} - I)\,\omega_2 + I_{23}\omega_3 &= 0 \\ I_{31}\omega_1 + I_{32}\omega_2 + (I_{33} - I)\,\omega_3 &= 0 \end{aligned} \tag{11.18}$$

This set of equations only has a non-trivial solution if the determinant of the coefficients vanish. This requires that

$$\begin{vmatrix} I_{11}-I & I_{12} & I_{13} \\ I_{21} & I_{22}-I & I_{23} \\ I_{31} & I_{32} & I_{33}-I \end{vmatrix} = 0 \tag{11.19}$$

This requirement leads to three possible values of I. Each of these corresponds to the moment of inertia about one of the principal aces.

11.3.1 Example: Problem 11.13

A three-particle system consists of masses m_i and coordinates (x_1, x_2, x_3) as follows:

$$\begin{aligned} m_1 &= 3m \qquad (b, 0, b) \\ m_2 &= 4m \qquad (b, b, -b) \\ m_3 &= 2m \qquad (-b, b, 0) \end{aligned} \tag{11.20}$$

Find the inertia tensor, the principal axes, and the principal moments of inertia.

We get the elements of the inertia tensor from Eq. 11.13a:

$$\begin{aligned} I_{11} &= \sum_\alpha m_\alpha \left(x_{\alpha,2}^2 + x_{\alpha,3}^2\right) = \\ &= 3m\left(b^2\right) + 4m\left(2b^2\right) + 2m\left(b^2\right) = 13mb^2 \end{aligned} \tag{11.21}$$

Likewise $I_{22} = 16mb^2$ and $I_{33} = 15mb^2$.

$$\begin{aligned} I_{12} &= I_{21} = -\sum_\alpha m_\alpha x_{\alpha,1} x_{\alpha,2} \\ &= -4m\left(b^2\right) - 2m\left(-b^2\right) = -2mb^2 \end{aligned} \tag{11.22}$$

Likewise $I_{13} = I_{31} = mb^2$ and $I_{23} = I_{32} = 4mb^2$. Thus, the inertia tensor is

$$\boxed{\{I\} = mb^2 \begin{bmatrix} 13 & -2 & 1 \\ -2 & 16 & 4 \\ 1 & 4 & 15 \end{bmatrix}} \tag{11.23}$$

The principal moments of inertia can be determined by solving the following equation

$$mb^2 \begin{bmatrix} 13-\lambda & -2 & 1 \\ -2 & 16-\lambda & 4 \\ 1 & 4 & 15-\lambda \end{bmatrix} = 0 \tag{11.24}$$

Expanding the determinant gives a cubic equation in λ:

$$\lambda^3 - 44\lambda^2 + 622\lambda - 2820 = 0 \tag{11.25}$$

Solving this equation numerically we obtain the following values for λ

$$\begin{aligned} \lambda_1 &= 10.00 \\ \lambda_2 &= 14.35 \\ \lambda_3 &= 19.65 \end{aligned} \tag{11.26}$$

The principle moments of inertia are equal to

$$\boxed{\begin{aligned} I_1 &= 10mb^2 \\ I_2 &= 14.35mb^2 \\ I_3 &= 19.65mb^2 \end{aligned}} \tag{11.27}$$

To find the principal axes, we substitute the different values of λ into Eq. 11.18 (see examples 11.3 and 11.5 in the textbook):

$$\begin{aligned} &(13-\lambda_i)\,\omega_{1i} - 2\omega_{2i} + \omega_{3i} = 0 \\ &-2\omega_{1i} + (16-\lambda_i)\,\omega_{2i} + 4\omega_{3i} = 0 \\ &\omega_{1i} + 4\omega_{2i} + (15-\lambda_i)\,\omega_{3i} = 0 \end{aligned} \tag{11.28}$$

For $i = 1$, we have $\lambda_1 = 10$:

$$\begin{aligned} &3\omega_{11} - 2\omega_{21} + \omega_{31} = 0 \\ &-2\omega_{11} + 6\omega_{21} + 4\omega_{31} = 0 \\ &\omega_{11} + 4\omega_{21} + 5\omega_{31} = 0 \end{aligned} \tag{11.29}$$

Solving the first equation for ω_{31} and substituting into the second equation shows that

$$\omega_{11} = \omega_{21} \tag{11.30}$$

Substituting this relation into the third equation shows that

$$\omega_{31} = -\omega_{21} \tag{11.31}$$

or

$$\omega_{11} : \omega_{21} : \omega_{31} = 1 : 1 : -1 \tag{11.32}$$

So, the principal axis associated with I_1 is

$$\boxed{\frac{1}{\sqrt{3}}\left(\hat{x} + \hat{y} - \hat{z}\right)} \tag{11.33}$$

Proceeding in the same way gives the other two principal axes:

$$\boxed{\begin{aligned} i = 2 &: -.81\hat{x} + .29\hat{y} - .52\hat{z} \\ i = 3 &: -.14\hat{x} + .77\hat{y} + .63\hat{z} \end{aligned}} \tag{11.34}$$

We note that the principal axes are mutually orthogonal, as they must be.

Our observation in problem 11.13 that the principal vectors are orthogonal is true in general. We can prove this in the following manner. For the m^{th} principal moment the following relations must hold:

$$L_{im} = I_m \omega_{im} \tag{11.35}$$

$$L_{im} = \sum_k I_{ik} \omega_{km} \tag{11.36}$$

Combining these two equations we obtain

$$\sum_k I_{ik}\omega_{km} = I_m\omega_{im} \tag{11.37}$$

Now multiply both sides of this equation by ω_{in} and sum over i:

$$\sum_{i,k} I_{ik}\omega_{km}\omega_{in} = \sum_i I_m\omega_{im}\omega_{in} \tag{11.38}$$

A similar relation can be obtained for the nth principal moment, multiplied by ω_{km} and summed over k:

$$\sum_{i,k} I_{ki}\omega_{in}\omega_{km} = \sum_k I_n\omega_{km}\omega_{kn} \tag{11.39}$$

If we subtract Eq. 11.39from Eq. 11.40 we obtain the following result:

$$\begin{aligned}\textstyle\sum_{i,k} I_{ik}\omega_{km}\omega_{in} - \sum_{i,k} I_{ki}\omega_{in}\omega_{km} &= \textstyle\sum_{i,k}\left(I_{ik} - I_{ki}\right)\omega_{in}\omega_{km} = 0 = \\ &= \textstyle\sum_i I_m\omega_{im}\omega_{in} - \sum_i I_n\omega_{im}\omega_{in} = \left(I_m - I_n\right)\sum_i \omega_{im}\omega_{in}\end{aligned} \tag{11.40}$$

Assuming that the principal momenta are distinct, the previous equation can only be correct if

$$\sum_i \omega_{im}\omega_{in} = \vec{\omega}_m \cdot \vec{\omega}_n = 0 \tag{11.41}$$

which shows the principal axes are orthogonal.

11.4 Transformations of the Inertia Tensor

In our discussion so far we have assumed that the origin of the rotating reference frame coincidence with the center of mass of the rigid object. In this Section we will examine what will change if we do not make this assumption.

Consider the two coordinate systems shown in Fig. 11.2. One reference frame, the x frame, has its origin O coincide with the center of mass of the rigid object; the second reference frame, the X frame, has an origin Q that is displaced with respect to the center of mass of the rigid object. The inertia tensor J_{ij} in reference frame X is defined in the same way as it was defined previously:

$$J_{ij} = \sum_\alpha m_\alpha\left(\delta_{ij}\sum_k {X_{\alpha,k}}^2 - X_{\alpha,i}X_{\alpha,j}\right) \tag{11.42}$$

The coordinates in the X frame are related to the coordinates in the x frame in the following way:

$$X_i = a_i + x_i \tag{11.43}$$

Using this relation, we can express the inertia tensor in reference frame X in terms of the coordinates in reference frame x:

$$\begin{aligned}J_{ij} &= \textstyle\sum_\alpha m_\alpha\left(\delta_{ij}\sum_k\left(a_k + x_{\alpha,k}\right)^2 - \left(a_i + x_{\alpha,i}\right)\left(a_j + x_{\alpha,j}\right)\right) = \\ &= \textstyle\sum_\alpha m_\alpha\left(\delta_{ij}\sum_k {x_{\alpha,k}}^2 - x_{\alpha,i}x_{\alpha,j}\right) + \sum_\alpha m_\alpha\left(\delta_{ij}\sum_k\left({a_k}^2 + 2a_kx_{\alpha,k}\right) - \left(a_ia_j + a_ix_{\alpha,j} + a_jx_{\alpha,i}\right)\right) = \\ &= \textstyle I_{ij} + \sum_\alpha m_\alpha\left(\delta_{ij}\sum_k {a_k}^2 - a_ia_j\right) + \sum_\alpha m_\alpha\left(2\delta_{ij}\sum_k\left(a_kx_{\alpha,k}\right) - a_ix_{\alpha,j} - a_jx_{\alpha,i}\right)\end{aligned} \tag{11.44}$$

The last term on the right-hand side is equal to 0 since the origin of the coordinate system x coincides with the center of mass of the object:

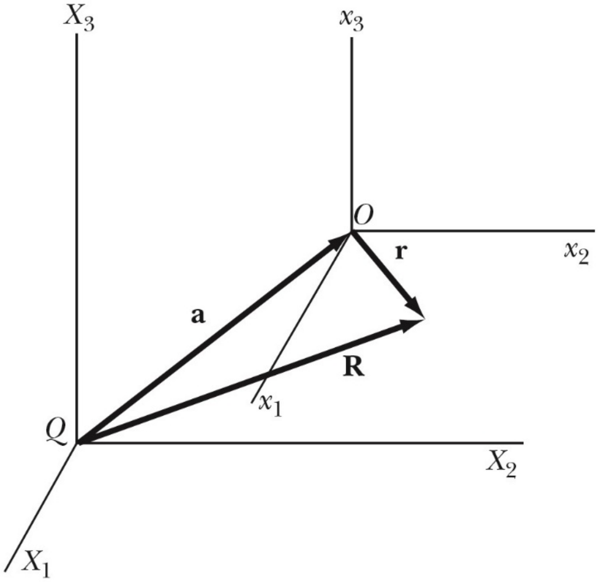


Figure 11.2: Two coordinate systems used to describe our rigid body.

$$\sum_\alpha m_\alpha x_{\alpha,k} = M r_{cm,k} = 0 \tag{11.45}$$

The relation between the inertia tensor in reference frame X and the inertia tensor in reference frame x is thus given by

$$J_{ij} = I_{ij} + \sum_\alpha m_\alpha \left(\delta_{ij} \sum_k {a_k}^2 - a_i a_j \right) = I_{ij} + M\left(\delta_{ij} a^2 - a_i a_j\right) \tag{11.46}$$

This relation is called the **Steiner's parallel-axis theorem** and is one example of how coordinate transformations affect the inertia tensor.
The transformation discussed so far is a simple translation. Other important transformations are rotations. In Chapter 1 we discussed many examples of rotations, and determined that the most general way to express rotations is by using the rotation matrix λ:

$$x_i = \sum_j \lambda_{ji} x'_j \tag{11.47}$$

Since this transformation rule is valid for vectors in general, the same rule can be used to describe the transformation of the angular momentum and angular velocity vectors:

$$L_i = \sum_j \lambda_{ji} L'_j \tag{11.48}$$

$$\omega_i = \sum_j \lambda_{ji} \omega'_j \tag{11.49}$$

In order to determine the relation between the inertia tensor in the two coordinate frames, we use the fact that the angular momentum is the product of the inertia tensor and the angular velocity, in both frames:

$$L_k = \sum_l I_{kl} \omega_l \tag{11.50}$$

and

$$L'_k = \sum_l I'_{kl}\omega'{}_l \tag{11.51}$$

In order to relate the inertia tensors, we use the coordinate transformations for L and ω:

$$\sum_j \lambda_{jk} L'_j = \sum_l I_{kl} \sum_m \lambda_{ml}\omega'_m \tag{11.52}$$

This equation can be simplified if we multiply each side by λ_k and sum over k:

$$\begin{aligned} \textstyle\sum_k \lambda_{ik}\left(\sum_j \lambda_{jk} L'_j\right) &= \textstyle\sum_j \left\{\sum_k \left(\lambda_{jk}\lambda_{ik}\right) L'_j\right\} = \sum_j \left\{\delta_{ji} L'_j\right\} = L'_i = \\ &= \textstyle\sum_k \lambda_{ik}\left(\sum_l I_{kl} \sum_m \lambda_{ml}\omega'_m\right) = \sum_m \left\{\sum_{k,l} \lambda_{ik}\lambda_{ml} I_{kl}\right\}\omega'_m \end{aligned} \tag{11.53}$$

where we have used the orthogonal properties of the rotation matrix. Using the relation between the angular momentum and the angular velocity in the rotated coordinate frame we see that the inertia tensors in the two coordinate frames are related as follows:

$$I'{}_{im} = \sum_{k,l} \lambda_{ik}\lambda_{ml} I_{kl} = \sum_{k,l} \lambda_{ik} I_{kl} \lambda^t{}_{lm} \tag{11.54}$$

where λ^t is the transposed matrix. In tensor notation we can rewrite this relation as

$$\{I'\} = \{\lambda\}\{I\}\left\{\lambda^t\right\} \tag{11.55}$$

It turns out that for any inertia tensor we can find a rotation such that the inertia tensor in the rotated frame is a diagonal matrix (all non-diagonal elements are equal to 0).
We thus have seen two different approaches to diagonalize the inertia tensor: 1) find the principal axes of inertia, and 2) find the proper rotation matrix.

11.4.1 Example: Problem 11.16

Consider the following inertia tensor:

$$\{I\} = \left\{ \begin{array}{ccc} \frac{1}{2}(A+B) & \frac{1}{2}(A-B)) & 0 \\ \frac{1}{2}(A-B)) & \frac{1}{2}(A+B) & 0 \\ 0 & 0 & C \end{array} \right\} \tag{11.56}$$

Perform a rotation of the coordinate system by an angle θ about the x_3 axis. Evaluate the transformed tensor elements, and show that the choice $\theta = \pi/4$ renders the inertia tensor diagonal with elements A, B, and C.

The rotation matrix is

$$\{\lambda\} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{11.57}$$

The moment of inertia tensor transforms according to

$$\{I'\} = \{\lambda\}\{I\}\{\lambda'\} \tag{11.58}$$

That is

$$\{I'\} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2}(A+B) & \frac{1}{2}(A-B) & 0 \\ \frac{1}{2}(A-B) & \frac{1}{2}(A+B) & 0 \\ 0 & 0 & C \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2}(A+B)\cos\theta + \frac{1}{2}(A-B)\sin\theta & -\frac{1}{2}(A+B)\sin\theta + \frac{1}{2}(A-B)\cos\theta & 0 \\ \frac{1}{2}(A-B)\cos\theta + \frac{1}{2}(A+B)\sin\theta & -\frac{1}{2}(A-B)\sin\theta + \frac{1}{2}(A+B)\cos\theta & 0 \\ 0 & 0 & C \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{1}{2}(A+B)\cos^2\theta + (A-B)\cos\theta\sin\theta + \frac{1}{2}(A+B)\sin^2\theta & \frac{1}{2}(A-B)\cos^2\theta - \frac{1}{2}(A-B)\sin^2\theta & 0 \\ -\frac{1}{2}(A-B)\sin^2\theta + \frac{1}{2}(A-B)\cos^2\theta & \frac{1}{2}(A+B)\sin^2\theta - (A-B)\sin\theta\cos\theta + \frac{1}{2}(A+B)\cos^2\theta & 0 \\ 0 & 0 & C \end{bmatrix} \tag{11.59}$$

or

$$\{I'\} = \begin{bmatrix} \frac{1}{2}(A+B) + (A-B)\cos\theta\sin\theta & \frac{1}{2}(A-B)\cos^2\theta - \frac{1}{2}(A-B)\sin^2\theta & 0 \\ -\frac{1}{2}(A-B)\sin^2\theta + \frac{1}{2}(A-B)\cos^2\theta & \frac{1}{2}(A+B) - (A-B)\cos\theta\sin\theta & 0 \\ 0 & 0 & C \end{bmatrix} \tag{11.60}$$

If $\theta = \pi/4$, $\sin\theta = \cos\theta = 1/\sqrt{2}$. Then,

$$\boxed{\{I'\} = \begin{bmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{bmatrix}} \tag{11.61}$$

11.5 Euler Angles

Any rotation between different coordinate systems can be expressed in terms of three successive rotations around the coordinate axes. When we consider the transformation from the fixed coordinate system x' to the body coordinate system x, we call the three angles the Euler angles ϕ, θ, and ψ (see Fig. 11.3). The total transformation matrix is the product of the individual transformations (note the order of the transformations):

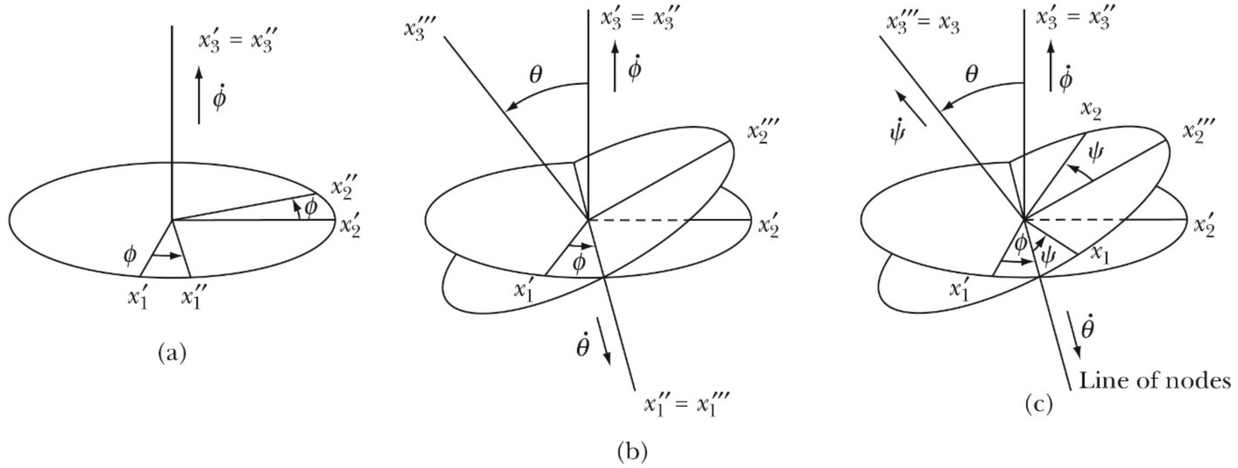


Figure 11.3: The Euler angles used to transform the fixed coordinate system x' into the body coordinate system x.

$$\lambda = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} = \tag{11.62}$$

$$= \begin{pmatrix} \cos\psi\cos\phi - \cos\theta\sin\phi\sin\psi & \cos\psi\sin\phi + \cos\theta\cos\phi\sin\psi & \sin\psi\sin\theta \\ -\sin\psi\cos\phi - \cos\theta\sin\phi\cos\psi & -\sin\psi\sin\phi + \cos\theta\cos\phi\cos\psi & \cos\psi\sin\theta \\ \sin\theta\sin\phi & -\sin\theta\cos\phi & \cos\theta \end{pmatrix}$$

With each of the three rotations we can associate an angular velocity ω. To express the angular velocity in the body coordinate system, we can use Fig. 11.3c.

1. ω_ϕ: Figure 3c shows that the angular velocity ω_ϕ is directed in the x_2''' - x_3''' plane. Its projection along the x_3''' axis, which is also the x_3 axis, is equal to

   $$\dot{\phi}_3 = \dot{\phi}\cos\theta \tag{11.63}$$

   The projection along the x_2''' axis is equal to

   $$\dot{\phi}_{2'''} = \dot{\phi}\sin\theta \tag{11.64}$$

   Fig. 11.3c shows that when we project this projection along the x_1 and x_2 axes we obtain the following components in the body coordinate system:

   $$\dot{\phi}_1 = \dot{\phi}_{2'''}\sin\psi = \dot{\phi}\sin\theta\sin\psi \tag{11.65}$$

   $$\dot{\phi}_2 = \dot{\phi}_{2'''}\cos\psi = \dot{\phi}\sin\theta\cos\psi \tag{11.66}$$

2. ω_θ: Fig. 11.3c shows that the angular velocity ω_θ is directed in the x_1''' - x_2''' plane. Its projection along the x_3''' axis, which is also the x_3 axis, is equal to 0.

$$\dot{\theta}_3 = 0 \tag{11.67}$$

Fig. 11.3c shows that when we project ω_θ along the x_1 and x_2 axes we obtain the following components in the body coordinate system:

$$\dot{\theta}_1 = \dot{\theta}\cos\psi \tag{11.68}$$

$$\dot{\theta}_2 = -\dot{\theta}\sin\psi \tag{11.69}$$

3. ω_ψ: Fig. 11.3c shows that the angular velocity ω_ψ is directed along the x_3''' axis, which is also the x_3 axis. The components along the other body axes are 0. Thus:

$$\dot{\psi}_1 = 0 \tag{11.70}$$

$$\dot{\psi}_2 = 0 \tag{11.71}$$

$$\dot{\psi}_3 = \dot{\psi} \tag{11.72}$$

The angular velocity, in the body frame, is thus equal to

$$\vec{\omega} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} = \begin{pmatrix} \dot{\phi}_1 + \dot{\theta}_1 + \dot{\psi}_1 \\ \dot{\phi}_2 + \dot{\theta}_2 + \dot{\psi}_2 \\ \dot{\phi}_3 + \dot{\theta}_3 + \dot{\psi}_3 \end{pmatrix} = \begin{pmatrix} \dot{\phi}\sin\theta\sin\psi + \dot{\theta}\cos\psi \\ \dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi \\ \dot{\phi}\cos\theta + \dot{\psi} \end{pmatrix} \tag{11.73}$$

11.6 The Force-Free Euler Equations

Let us assume for the moment that the coordinate axes correspond to the principal axes of the body. In that case, we can write the kinetic energy of the body in the following manner:

$$T = \frac{1}{2}\sum_i I_i {\omega_i}^2 \tag{11.74}$$

where I_i are the principal moments of the rigid body. If for now we consider that the rigid object is carrying out a force-free motion ($U = 0$) then the Lagrangian L will be equal to the kinetic energy T. The motion of the object can be described in terms of the Euler angles, which can serve as the generalized coordinates of the motion. Consider the three equations of motion for the three generalized coordinates:

1. The Euler angle ϕ: Lagrange's equation for the coordinate ϕ is

$$0 = \frac{\partial L}{\partial \phi} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} = \frac{\partial T}{\partial \phi} - \frac{d}{dt}\frac{\partial T}{\partial \dot{\phi}} = \sum_i \frac{\partial T}{\partial \omega_i}\frac{\partial \omega_i}{\partial \phi} - \frac{d}{dt}\sum_i \frac{\partial T}{\partial \omega_i}\frac{\partial \omega_i}{\partial \dot{\phi}} = \sum_i I_i \omega_i \frac{\partial \omega_i}{\partial \phi} - \frac{d}{dt}\sum_i I_i \omega_i \frac{\partial \omega_i}{\partial \dot{\phi}} = 0 \tag{11.75}$$

Differentiating the angular velocity with respect to the coordinate ϕ we find

$$\begin{aligned} \frac{\partial\omega_1}{\partial\phi} &= 0 \qquad & \frac{\partial\omega_1}{\partial\dot{\phi}} &= \sin\theta\sin\psi \\ \frac{\partial\omega_2}{\partial\phi} &= 0 \qquad & \frac{\partial\omega_2}{\partial\dot{\phi}} &= \sin\theta\cos\psi \\ \frac{\partial\omega_3}{\partial\phi} &= 0 \qquad & \frac{\partial\omega_3}{\partial\dot{\phi}} &= \cos\theta \end{aligned} \tag{11.76}$$

and Lagrange's equation of motion becomes

$$\frac{d}{dt}\left\{I_1\omega_1\sin\theta\sin\psi + I_2\omega_2\sin\theta\cos\psi + I_3\omega_3\cos\theta\right\} = 0 \tag{11.77}$$

2. The Euler angle θ: Lagrange's equation for the coordinate θ is

$$0 = \frac{\partial T}{\partial\theta} - \frac{d}{dt}\frac{\partial T}{\partial\dot{\theta}} = \sum_i \frac{\partial T}{\partial\omega_i}\frac{\partial\omega_i}{\partial\theta} - \frac{d}{dt}\sum_i \frac{\partial T}{\partial\omega_i}\frac{\partial\omega_i}{\partial\dot{\theta}} = \sum_i I_i\omega_i\frac{\partial\omega_i}{\partial\theta} - \frac{d}{dt}\sum_i I_i\omega_i\frac{\partial\omega_i}{\partial\dot{\theta}} = 0 \tag{11.78}$$

Differentiating the angular velocity with respect to the coordinate θ we find

$$\begin{aligned} \frac{\partial\omega_1}{\partial\theta} &= \dot{\phi}\cos\theta\sin\psi \qquad & \frac{\partial\omega_1}{\partial\dot{\theta}} &= \cos\psi \\ \frac{\partial\omega_2}{\partial\theta} &= \dot{\phi}\cos\theta\cos\psi \qquad & \frac{\partial\omega_2}{\partial\dot{\theta}} &= -\sin\psi \\ \frac{\partial\omega_3}{\partial\theta} &= -\dot{\phi}\sin\theta \qquad & \frac{\partial\omega_3}{\partial\dot{\theta}} &= 0 \end{aligned} \tag{11.79}$$

and Lagrange's equation of motion becomes

$$\dot{\phi}\left(\left\{I_1\omega_1\sin\psi + I_2\omega_2\cos\psi\right\}\cos\theta - I_3\omega_3\sin\theta\right) - \frac{d}{dt}\left\{I_1\omega_1\cos\psi - I_2\omega_2\sin\psi\right\} = 0 \tag{11.80}$$

3. The Euler angle ψ: Lagrange's equation for the coordinate ψ is

$$0 = \frac{\partial T}{\partial\psi} - \frac{d}{dt}\frac{\partial T}{\partial\dot{\psi}} = \sum_i \frac{\partial T}{\partial\omega_i}\frac{\partial\omega_i}{\partial\psi} - \frac{d}{dt}\sum_i \frac{\partial T}{\partial\omega_i}\frac{\partial\omega_i}{\partial\dot{\psi}} = \sum_i I_i\omega_i\frac{\partial\omega_i}{\partial\psi} - \frac{d}{dt}\sum_i I_i\omega_i\frac{\partial\omega_i}{\partial\dot{\psi}} = 0 \tag{11.81}$$

Differentiating the angular velocity with respect to the coordinate y we find

$$\begin{aligned} \frac{\partial\omega_1}{\partial\psi} &= \dot{\phi}\sin\theta\cos\psi - \dot{\theta}\sin\psi = \omega_2 \qquad & \frac{\partial\omega_1}{\partial\dot{\psi}} &= 0 \\ \frac{\partial\omega_2}{\partial\psi} &= -\dot{\phi}\sin\theta\sin\psi - \dot{\theta}\cos\psi = -\omega_1 \qquad & \frac{\partial\omega_2}{\partial\dot{\psi}} &= 0 \\ \frac{\partial\omega_3}{\partial\psi} &= 0 \qquad & \frac{\partial\omega_3}{\partial\dot{\psi}} &= 1 \end{aligned} \tag{11.82}$$

and Lagrange's equation of motion becomes

$$I_1\omega_1\omega_2 - I_2\omega_2\omega_1 - \frac{d}{dt}\left\{I_3\omega_3\right\} = (I_1 - I_2)\,\omega_1\omega_2 - \frac{d}{dt}\left\{I_3\omega_3\right\} = (I_1 - I_2)\,\omega_1\omega_2 - I_3\dot{\omega}_3 = 0 \tag{11.83}$$

Of all three equations of motion, the last one is the only one to contain just the components of the angular velocity. Since our choice of the x_3 axis was arbitrary, we expect that similar relations should exist for the other two axes. The set of three equation we obtain in this way are called the **Euler equations**:

$$(I_1 - I_2)\,\omega_1\omega_2 - I_3\dot{\omega}_3 = 0 \tag{11.84}$$

$$(I_2 - I_3)\,\omega_2\omega_3 - I_1\dot{\omega}_1 = 0 \tag{11.85}$$

$$(I_3 - I_1)\,\omega_3\omega_1 - I_2\dot{\omega}_2 = 0 \tag{11.86}$$

As an example of how we use Euler's equations, consider a symmetric top. The top will have two different principal moments: $I_1 = I_2$ and I_3. In this case, the first Euler equations reduces to

$$I_3\dot{\omega}_3 = 0 \tag{11.87}$$

or

$$\omega_3\,(t) = \text{constant} = \omega_3 \tag{11.88}$$

The other two Euler equations can be rewritten as

$$\dot{\omega}_1 = -\left(\frac{I_3 - I_1}{I_1}\omega_3\right)\omega_2 = -\omega\omega_2 \tag{11.89}$$

$$\dot{\omega}_2 = \left(\frac{I_3 - I_1}{I_1}\omega_3\right)\omega_1 = \omega\omega_1 \tag{11.90}$$

where ω is

$$\omega = \left(\frac{I_3 - I_1}{I_1}\omega_3\right) \tag{11.91}$$

This set of equations has the following solution:

$$\omega_1\,(t) = A\cos\omega t \tag{11.92}$$

$$\omega_2\,(t) = A\sin\omega t \tag{11.93}$$

The magnitude of the angular velocity of the system is constant since

$$\left|\vec{\omega}\right| = \sqrt{\left({\omega_1}^2\,(t) + {\omega_2}^2\,(t) + {\omega_3}^2\right)} = \sqrt{A^2 + {\omega_3}^2} \tag{11.94}$$

The angular velocity vector traces out a cone in the body frame (it precesses around the x_3 axis - see Fig. 11.4). The rate with which the angular velocity vector precesses around the x_3 axis is determined by the value of ω. When the principal moment I_3 and the principal moment I_1 are similar, ω becomes very small.

Since we have assumed that there are no external forces and torques acting on the system, the angular momentum of the system will be constant in the fixed reference frame. If the angular momentum is initially pointing along the x'_3 axis it will continue to point along this axis (see Fig. 11.5). Since there are no external forces and torques acting on the system, the rotation kinetic energy of the system must be constant. Thus

$$T_{rot} = \frac{1}{2}\vec{\omega}\cdot\vec{L} \tag{11.95}$$

Since the angle between the angular velocity vector and the angular momentum vector must be constant, the angular velocity vector must trace out a space cone around the x'_3 (see Fig. 11.5).

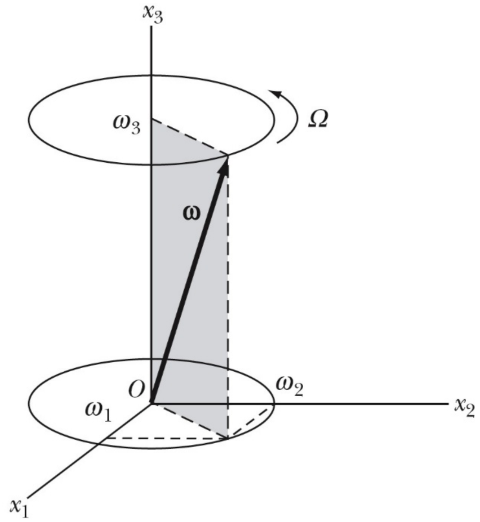


Figure 11.4: The angular velocity of a force-free symmetric top, precessing around the x_3 axis in the body frame.

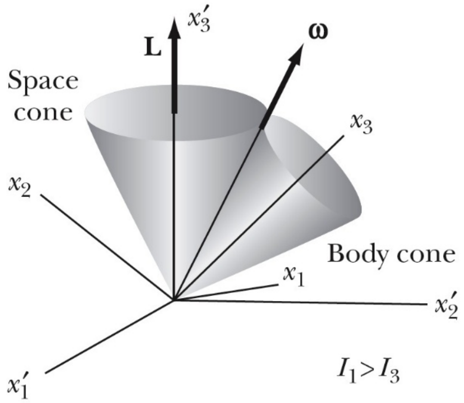


Figure 11.5: The angular velocity of a force-free symmetric top, tracing out a space-cone around the x_3' axis in the body frame.

11.6.1 Example: Problem 11.27

A symmetric body moves without the influence of forces or torques. Let x_3 be the symmetry axis of the body and L be along x_3'. The angle between the angular velocity vector and x_3 is α. Let ω and L initially be in the x_2-x_3 plane. What is the angular velocity of the symmetry axis about L in terms of I_1, I_3, ω, and α?

The coordinate system we will use to solve this problem is shown in Fig. 11.6. The initial conditions are given by

$$L_1 = 0 = I_1\omega_1 \tag{11.96}$$

$$L_2 = L\sin\theta = I_1\omega_2 = I_1\omega\sin\alpha \tag{11.97}$$

$$L_3 = L\cos\theta = I_3\omega_3 = I_3\omega\cos\alpha \tag{11.98}$$

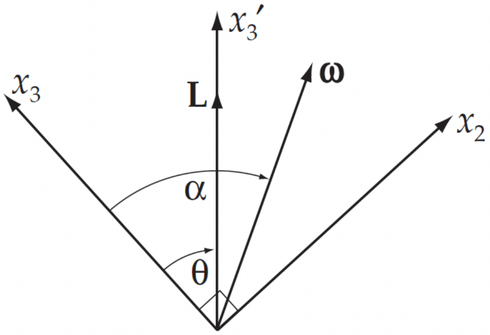


Figure 11.6: Coordinate system used to solve Problem 11.27.

Thus

$$\tan\theta = \frac{L_2}{L_3} = \frac{I_1}{I_3}\tan\alpha \tag{11.99}$$

Use Eq. (11.102) in the textbook:

$$\omega_3 = \dot{\phi}\cos\theta + \dot{\psi} \tag{11.100}$$

Since $\omega_3 = \omega\cos\alpha$, we have

$$\dot{\phi}\cos\theta = \omega\cos\alpha - \dot{\psi} \tag{11.101}$$

Using Eq. (11.131) in the textbook

$$\dot{\psi} = -\Omega = -\frac{I_3 - I_1}{I_1}\omega_3 \tag{11.102}$$

Eq. (11.101) becomes

$$\dot{\phi}\cos\theta = \frac{I_3}{I_1}\omega\cos\alpha \tag{11.103}$$

We can use Eq. (11.99) to express $\cos\theta$ in terms of $\tan\alpha$

$$\cos\theta = \frac{I_3}{\left[I_3^2 + I_1^2\tan^2\alpha\right]^{1/2}} \tag{11.104}$$

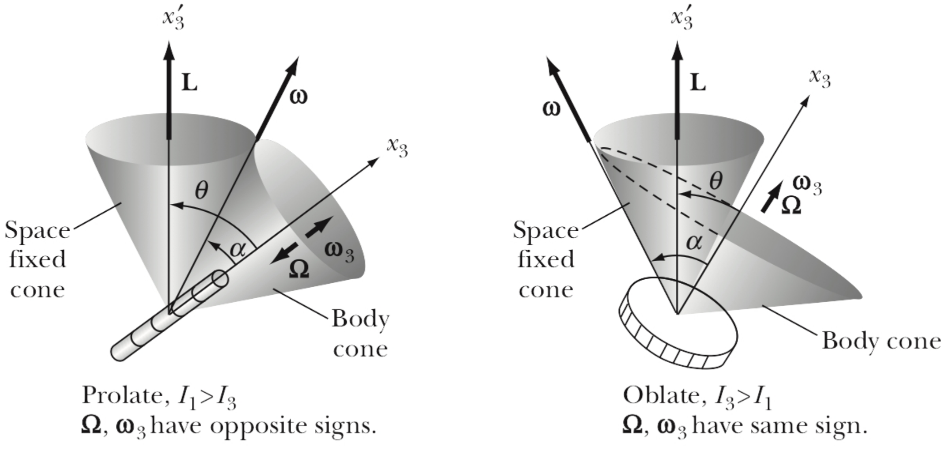
Substituting this expression for $\cos\theta$ into (11.103) gives

$$\dot{\phi} = \frac{\omega}{I_1}\sqrt{I_1^2\sin^2\alpha + I_3^2\cos^2\alpha} \tag{11.105}$$

Equation 11.99 shows that

1. if $I_1 > I_3$ (prolate body), $\theta > \alpha$
2. if $I_1 < I_3$ (oblate body), $\theta < \alpha$

The space and body cones for these two different scenarios are shown in Fig. 11.7.



(a) (b)

Figure 11.7: Space and body cones for prolate and oblate bodies

11.7 The Euler Equations in a Force Field

When the external forces and torques acting on the system are not equal to 0, we can not use the method we have used in the previous section to obtain expressions for the angular velocity and acceleration. The procedure used in the previous section relied on the fact that the potential energy U is 0 in a force-free environment, and therefore, the Lagrangian L is equal to the kinetic energy T.

When the external forces and torques are not equal to 0, the angular momentum of the system is not conserved:

$$\left(\frac{d\vec{L}}{dt}\right)_{\text{fixed}} = \vec{N} \tag{11.106}$$

Note that this relation only holds in the fixed reference frame since this is the only good inertial reference frame. In Chapter 10 we looked at the relation between parameters specified in the fixed reference frame compared to parameters specified in the rotating reference frame, and we can use this relation to correlate the rate of change of the angular momentum vector in the fixed reference frame with the rate of change of the angular momentum vector in the rotating reference frame:

$$\vec{N} = \left(\frac{d\vec{L}}{dt}\right)_{\text{fixed}} = \left(\frac{d\vec{L}}{dt}\right)_{\text{rotating}} + \vec{\omega} \times \vec{L} \tag{11.107}$$

This relation can be used to generate three separate relations by projecting the vectors along the three body axes:

$$N_1 = \frac{dL_1}{dt} + \left(\vec{\omega} \times \vec{L}\right)_1 = \frac{dL_1}{dt} + (\omega_2 L_3 - \omega_3 L_2) = I_1 \dot{\omega}_1 - (I_2 - I_3)\,\omega_2 \omega_3 \tag{11.108}$$

$$N_2 = \frac{dL_2}{dt} + \left(\vec{\omega} \times \vec{L}\right)_2 = \frac{dL_2}{dt} + (\omega_3 L_1 - \omega_1 L_3) = I_2 \dot{\omega}_2 - (I_3 - I_1)\,\omega_3 \omega_1 \tag{11.109}$$

$$N_3 = \frac{dL_3}{dt} + \left(\vec{\omega} \times \vec{L}\right)_3 = \frac{dL_3}{dt} + (\omega_1 L_2 - \omega_2 L_1) = I_3 \dot{\omega}_3 - (I_1 - I_2)\,\omega_1 \omega_2 \tag{11.110}$$

These equations are the Euler equations for the motion of the rigid body in a force field. In the absence of a torque, these equations reduce to the force-free Euler equations.

11.7.1 Example: Motion of a Symmetric Top with One Point Fixed

In order to describe the motion of a top, which has its tip fixed, we use two coordinate systems whose origins coincide (see Fig. 11.8). Since the origins coincide, the transformation between coordinate systems can be described in terms of the Euler angles, and the equations of motion will be the Euler equations:

$$(I_1 - I_2)\,\omega_1\omega_2 - I_3\dot{\omega}_3 = 0 \tag{11.111}$$

$$(I_2 - I_3)\,\omega_2\omega_3 - I_1\dot{\omega}_1 = N_1 \tag{11.112}$$

$$(I_3 - I_1)\,\omega_3\omega_1 - I_2\dot{\omega}_2 = N_2 \tag{11.113}$$

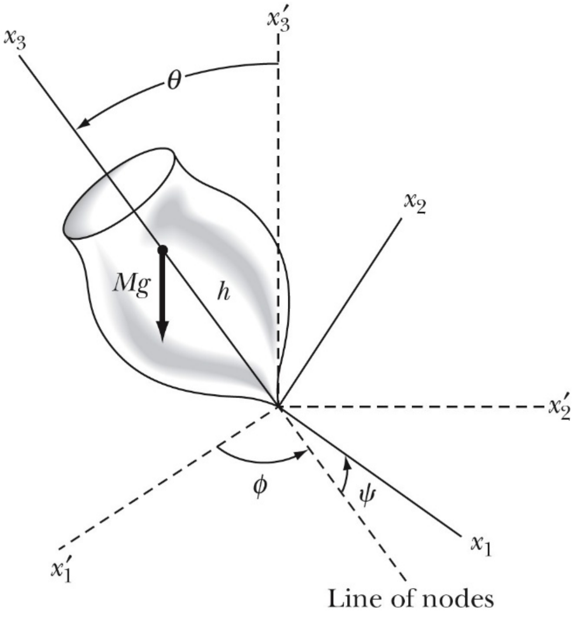


Figure 11.8: Spinning top with fixed tip.

Since the top is symmetric around the x_3 axis, it principal moments of inertia with respect to the x_1 and x_2 axes are identical. The Euler equations now become

$$I_3\dot{\omega}_3 = 0 \tag{11.114}$$

$$(I_1 - I_3)\,\omega_2\omega_3 - I_1\dot{\omega}_1 = N_1 \tag{11.115}$$

$$-(I_1 - I_3)\,\omega_3\omega_1 - I_1\dot{\omega}_2 = N_2 \tag{11.116}$$

The first equation immediately tells us that

$$\omega_3 = \text{constant} \tag{11.117}$$

The motion of the top is often described in terms of the motion of its rotating axes. The kinetic energy of the system is equal to

$$T = \frac{1}{2}\sum_i I_i{\omega_i}^2 = \frac{1}{2}I_1\left({\omega_1}^2 + {\omega_2}^2\right) + \frac{1}{2}I_3{\omega_3}^2 = \frac{1}{2}I_1\left(\dot{\phi}^2\sin^2\theta + \dot{\theta}^2\right) + \frac{1}{2}I_3\left(\dot{\phi}\cos\theta + \dot{\psi}\right)^2 \tag{11.118}$$

The potential energy of the system, assuming the center of mass of the top is located a distance h from the tip, is equal to

$$U = Mgh\cos\theta \tag{11.119}$$

The Lagrangian is thus equal to

$$L = T - U = \frac{1}{2}I_1\left(\dot{\phi}^2\sin^2\theta + \dot{\theta}^2\right) + \frac{1}{2}I_3\left(\dot{\phi}\cos\theta + \dot{\psi}\right)^2 - Mgh\cos\theta \tag{11.120}$$

The Lagrangian does not depend on ϕ and y, and thus

$$\frac{\partial L}{\partial \phi} = 0 = \frac{d}{dt}\frac{\partial L}{\partial \dot{\phi}} \tag{11.121}$$

$$\frac{\partial L}{\partial \psi} = 0 = \frac{d}{dt}\frac{\partial L}{\partial \dot{\psi}} \tag{11.122}$$

We thus conclude that the angular momenta associated with the Euler angles ϕ and ψ are constant:

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = \dot{\phi}\left(I_1\sin^2\theta + I_3\cos^2\theta\right) + \dot{\psi}I_3\cos\theta = \text{constant} \tag{11.123}$$

$$p_\psi = \frac{\partial L}{\partial \dot{\psi}} = I_3\left(\dot{\psi} + \dot{\phi}\cos\theta\right) = I_3\omega_3 = \text{constant} \tag{11.124}$$

Expressing the momenta in terms of the Euler angles ϕ and ψ allows us to express the rate of change of these Euler angles in terms of the angular momenta:

$$\dot{\phi} = \frac{p_\phi - p_\psi\cos\theta}{I_1\sin^2\theta} \tag{11.125}$$

Since there are no non-conservative forces acting on the top, the total energy E of the system is conserved. Thus

$$E = T + U = \frac{1}{2}I_1\left(\dot{\phi}^2\sin^2\theta + \dot{\theta}^2\right) + \frac{1}{2}I_3\left(\dot{\phi}\cos\theta + \dot{\psi}\right)^2 + Mgh\cos\theta = \text{constant} \tag{11.126}$$

The total energy can be rewritten in terms of the angular momenta:

$$\begin{aligned} E &= \tfrac{1}{2}I_1\left(\dot{\phi}^2\sin^2\theta + \dot{\theta}^2\right) + \tfrac{1}{2}I_3\left(\dot{\phi}\cos\theta + \dot{\psi}\right)^2 + Mgh\cos\theta = \\ &= \tfrac{1}{2}I_1\left(\dot{\phi}^2\sin^2\theta + \dot{\theta}^2\right) + \tfrac{1}{2}I_3{\omega_3}^2 + Mgh\cos\theta = \\ &= \tfrac{1}{2}I_1\dot{\theta}^2 + \tfrac{1}{2}I_1\dot{\phi}^2\sin^2\theta + \tfrac{1}{2}I_3{\omega_3}^2 + Mgh\cos\theta = \\ &= \tfrac{1}{2}I_1\dot{\theta}^2 + \tfrac{1}{2}\tfrac{\left(p_\phi - p_\psi\cos\theta\right)^2}{I_1\sin^2\theta} + \tfrac{1}{2}I_3{\omega_3}^2 + Mgh\cos\theta \end{aligned} \tag{11.127}$$

Since the angular velocity with respect to the x_3 axis is constant, we can subtract it from the energy E to get the effective energy E' (note: this is equivalent to choosing the zero point of the energy scale). Thus

$$E' = E - \frac{1}{2}I_3{\omega_3}^2 = \frac{1}{2}I_1\dot{\theta}^2 + \frac{1}{2}\frac{\left(p_\phi - p_\psi\cos\theta\right)^2}{I_1\sin^2\theta} + Mgh\cos\theta = \text{constant} \tag{11.128}$$

The effective energy only depends on the angle θ and on $d\theta/dt$ since the angular momenta are constants. The manipulations we have carried out have reduced the three-dimensional problem to a one-dimensional problem. The first term in the effective energy is the kinetic energy associated with the rotation around the x_1 axis. The last two terms depend only on the angle θ and not on the angular velocity $d\theta/dt$. These terms are what we could call the effective potential energy, defined as

$$V(\theta) = \frac{1}{2}\frac{\left(p_\phi - p_\psi \cos\theta\right)^2}{I_1 \sin^2\theta} + Mgh\cos\theta \tag{11.129}$$

The effective potential becomes large when the angle approaches 0 and π. The angular dependence of the effective potential is shown in Fig. 11.9. If the total effective energy of the system is E_1', we expect the angle θ to vary between θ_1 and θ_2. We thus expect that the angle of inclination of the top will vary between these two extremes.

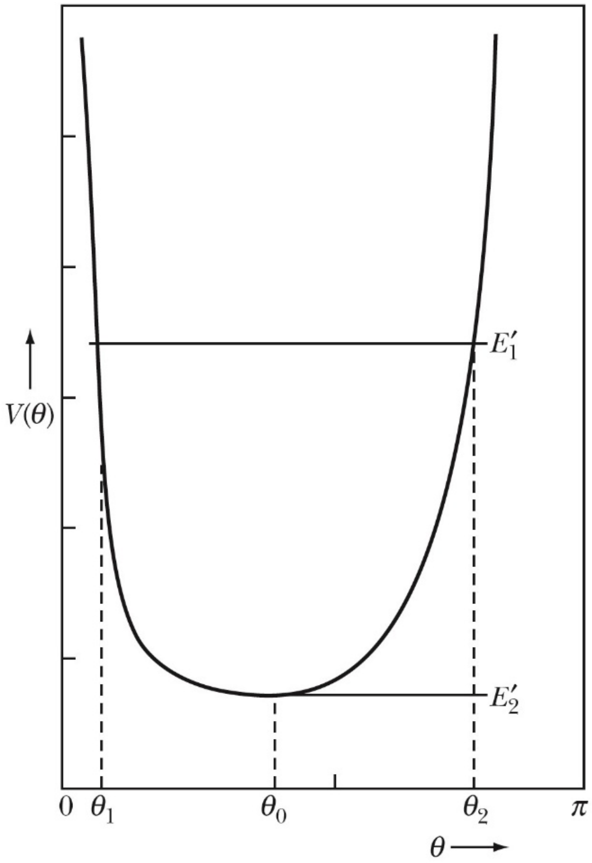


Figure 11.9: The effective potential of a rotating top.

The minimum effective energy that the system can have is E_2'. The corresponding angle can be found by requiring

$$\left.\frac{\partial V}{\partial \theta}\right|_{\theta=\theta_0} = 0 \tag{11.130}$$

This requirement can be rewritten as a quadratic equation of a parameter β, where β is defined as

$$\beta = p_\phi - p_\psi \cos\theta_0 \tag{11.131}$$

In general, there are two solutions to this quadratic equation. Since β is a real number, the solution must be real, and this requires that

$$1 - \frac{4MghI_1\cos\theta_0}{{p_\psi}^2} \geq 0 \tag{11.132}$$

This equation can be rewritten as

$$p_\psi{}^2 = (I_3\omega_3)^2 \geq 4MghI_1\cos\theta_0 \tag{11.133}$$

When we study a spinning top, the spin axis is oriented such that $\theta_0 < \pi/2$. The previous equation can then be rewritten as

$$I_3\omega_3 \geq \sqrt{4MghI_1\cos\theta_0} \tag{11.134}$$

or

$$\omega_3 \geq \frac{2}{I_3}\sqrt{MghI_1\cos\theta_0} \tag{11.135}$$

There is thus a minimum angular velocity the system must have in order to produce stable precession. The rate of precession can be found by calculating

$$\dot{\phi} = \frac{p_\phi - p_\psi\cos\theta_0}{I_1\sin^2\theta_0} = \frac{\beta}{I_1\sin^2\theta_0} \tag{11.136}$$

Since β has two possible values, we expect to see two different precession rates: one resulting in fast precession, and one resulting in slow precession.

When the angle of inclination is not equal to θ_0, the system will oscillate between two limiting values of θ. The precession rate will be a function of θ and can vary between positive and negative values, depending on the values of the angular momenta. The phenomenon is called nutation, and possible nutation patterns are shown in Fig. 11.10. The type of nutation depends on the initial conditions.

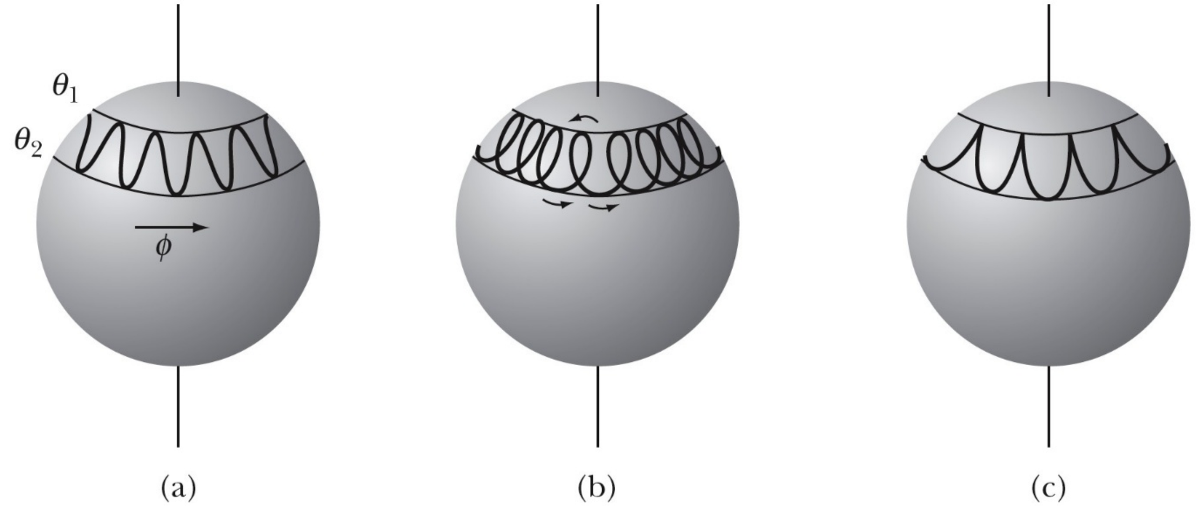


Figure 11.10: The nutation of a rotating top.

11.7.2 Example: Problem 11.30

Investigate the equation for the turning points of the nutational motion by setting $d\theta/dt = 0$ in the equation of the effective energy. Show that the resulting equation is cubic in $\cos\theta$ and has two real roots and one imaginary root.

If we set $\dot{\theta} = 0$ in the equation for the effective energy we obtain

$$E' = V(\theta) = \frac{\left(P_\phi - P_\psi\cos\theta\right)^2}{2I_{12}(1-\cos^2\theta)} + Mgh\cos\theta \tag{11.137}$$

Re-arranging, this equation can be written as

$$(2MghI_{12})\cos^3\theta - \left(2E'I_{12} + P_\psi^2\right)\cos^2\theta + 2\left(P_\phi P_\psi - MghI_{12}\right)\cos\theta + \left(2E'I_{12} - P_\phi^2\right) = 0 \tag{11.138}$$

which is cubic in $\cos\theta$. Equation 11.138 has the form shown in Fig. 11.11. Two of the roots occur in the region $-1 < \cos\theta < 1$. One root lies outside this range and is therefore imaginary.

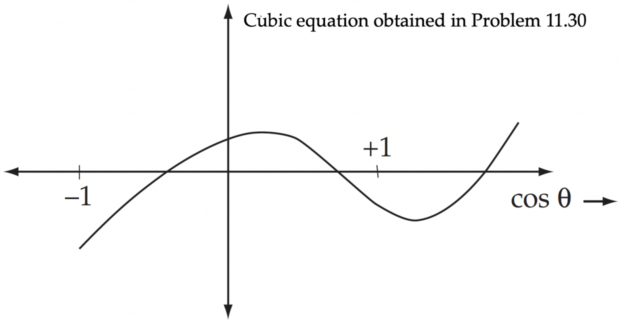


Figure 11.11: Equation 11.138, obtained in Problem 11.30, as function of $\cos\theta$.

11.8 Stability of Rigid-Body Rotations

The rotation of a rigid body is stable if the system, when perturbed from its equilibrium condition, carries out small oscillations about it. Consider we use the principal axes of rotation to describe the motion, and we choose these axes such that $I_3 > I_2 > I_1$. If the system rotates around the x_1 axis we can write the angular velocity vector as

$$\vec{\omega} = \omega_1 \hat{x}_1 \tag{11.139}$$

Consider what happens when we apply a small perturbation around the other two principal axes such that the angular velocity becomes

$$\vec{\omega} = \omega_1 \hat{x}_1 + \lambda \hat{x}_2 + \mu \hat{x}_3 \tag{11.140}$$

The corresponding Euler equations are

$$(I_1 - I_2)\,\omega_1\omega_2 - I_3\dot{\omega}_3 = (I_1 - I_2)\,\lambda\omega_1 - I_3\dot{\mu} = 0 \tag{11.141}$$

$$(I_2 - I_3)\,\omega_2\omega_3 - I_1\dot{\omega}_1 = (I_2 - I_3)\,\lambda\mu - I_1\dot{\omega}_1 = 0 \tag{11.142}$$

$$(I_3 - I_1)\,\omega_3\omega_1 - I_2\dot{\omega}_2 = (I_3 - I_1)\,\mu\omega_1 - I_2\dot{\lambda} = 0 \tag{11.143}$$

Since we are talking about small perturbations from the equilibrium state, $\lambda\mu$ will be small and can be set to 0. The second equation can thus be used to conclude that

$$\omega_1 = \text{constant} \tag{11.144}$$

The remaining equations can be rewritten as

$$\dot{\mu} = \left(\frac{I_1 - I_2}{I_3}\omega_1\right)\lambda \tag{11.145}$$

$$\dot{\lambda} = \left(\frac{I_3 - I_1}{I_2}\omega_1\right)\mu \tag{11.146}$$

The last equation can be differentiated to obtain

$$\ddot{\lambda} = \frac{d\dot{\lambda}}{dt} = \left(\frac{I_3 - I_1}{I_2}\omega_1\right)\frac{d\mu}{dt} = \left(\frac{I_3 - I_1}{I_2}\omega_1\right)\left(\frac{I_1 - I_2}{I_3}\omega_1\right)\lambda = -\left(\frac{(I_3 - I_1)(I_2 - I_1)}{I_2 I_3}\omega_1{}^2\right)\lambda \tag{11.147}$$

The term within the parenthesis is positive since we assumed that $I_3 > I_2 > I_1$. This differential equation has the following solution:

$$\lambda(t) = A_\lambda \cos(\Omega_1 t) + B_\lambda \sin(\Omega_1 t) \tag{11.148}$$

where

$$\omega_1 = \Omega_1 \sqrt{\frac{(I_3 - I_1)(I_2 - I_1)}{I_2 I_3}} \tag{11.149}$$

When we look at the perturbation around the x_3 axis we find the following differential equation

$$\ddot{\mu} = \frac{d\dot{\mu}}{dt} = \left(\frac{I_1 - I_2}{I_3}\omega_1\right)\frac{d\lambda}{dt} = \left(\frac{I_1 - I_2}{I_3}\omega_1\right)\left(\frac{I_3 - I_1}{I_2}\omega_1\right)\mu = -\left(\frac{(I_2 - I_1)(I_3 - I_1)}{I_2 I_3}\omega_1{}^2\right)\mu \tag{11.150}$$

The solution of the second-order differential equation is

$$\mu(t) = A_\mu \cos(\Omega_1 t) + B_\mu \sin(\Omega_1 t) \tag{11.151}$$

We see that the perturbations around the x_2 axis and the x_3 axis oscillate around the equilibrium values of $\lambda = \mu = 0$. We thus conclude that the rotation around the x_1 axis is stable.
Similar calculations can be done for rotations around the x_2 axis and the x_3 axis. The perturbation frequencies obtained in those cases are equal to

$$\Omega_2 = \omega_2 \sqrt{\frac{(I_1 - I_2)(I_3 - I_2)}{I_3 I_1}} \tag{11.152}$$

$$\Omega_3 = \omega_3 \sqrt{\frac{(I_2 - I_3)(I_1 - I_3)}{I_1 I_2}} \tag{11.153}$$

We see that Ω_2 is an imaginary number while Ω_3 is a real number. Thus, the rotation around the x_3 axis is stable, but the rotation around the x_2 axis is unstable.

11.8.1 Example: Problem 11.34

Consider a symmetrical rigid body rotating freely about its center of mass. A frictional torque $N_f = -b\omega$ acts to slow down the rotation. Find the component of the angular velocity along the symmetry axis as a function of time.

The Euler equation, which describes the rotation of an object about its symmetry axis, say the x axis, is

$$I_x \dot{\omega}_x - \left(I_y - I_z\right)\omega_y \omega_z = N_x \tag{11.154}$$

where $N_x = -b\omega_x$ is the component of torque along the x axis. Because the object is symmetric about the x axis, we have $I_y = I_z$, and the above equation becomes

$$I_x \frac{\mathrm{d}\omega_x}{\mathrm{d}t} = -b\omega_x \Rightarrow \omega_x = e^{-\frac{b}{I_x}t}\omega_{x0} \tag{11.155}$$