The Two-Body Problem with a $\frac{1}{r^2}$ Force

Arthur Jaffe

September 26, 2023

Contents

1 Background

The attractive force proportional to $\frac{1}{r^2}$ is called the "Kepler problem," as it is named after Kepler's, study of planetary motion, leading up to Newton. This inverse square force arises from gravity, and can be predicted as an approximation to Einstein's equations of relativity.

This attractive inverse square force law leads to three types of orbits, according to the value of the energy. They describe bound planets, unbound comets, and an intermediate case. The same equations, but with a repulsive force law, describe the scattering trajectories of electrically-charged particles, which interact without the influence of any external electric or magnetic field. In the repulsive case, there is only one type of orbit; these repulsive orbits were studied by Rutherford and led to the discovery of the nucleus.

In these notes we consider an inverse square force law between two bodies, idealized as particles. It is an interesting fact that the gravitational force of attraction of two uniform spheres of mass m and M is the same as the force of two point masses m and M, located at the centers of the spheres.

We start with the reduction of the two-body problem to a one-body problem. We then solve for the resulting orbits using two different methods.

i. The first method depends on finding sufficiently many conserved quantities: one scalar and two, three-component vectors. This gives the answer. This importance of the second method is related to a main focus of our course: conserved quantities are lead to symmetry of the solution. And symmetry is one of the most beautiful aspects of nature.

   The point of this course is to show that symmetry is both beautiful, and it provides a simplifying theme to understand nature. So we shine a light on this method.

ii. The second method is the traditional way. One solves a differential equation describing

Newton's law. This gives the same answer.

In these notes we will assume that the bodies we consider are point particles. One can justify this for an earth that is a uniform sphere. And although the earth is not a uniform sphere, that will be the approximation that we consider. Let us do the calculation for a uniform sphere of mass M and radius a, by calculating the gravitational potential outside the sphere. By the spherical symmetry of the problem, we can center the sphere at the origin and calculate the potential at a distance d along the positive, horizontal axis. In order to do this, we use polar coordinates in three dimensions.

1.1 The Gravitational Potential of a Uniform Sphere

Newton observed in problems of gravitation that one could replace the forces created by a uniform spherical body by the forces created by a point mass located at the center of the body. That justifies all the planetary calculations that we will do for equations that describe for point particles that representing the positions of the center of spherical objects. Of course the planets are not uniform spherical bodies, but then one can study and understand the deviations from that state. Here we consider that point particles of mass M and m separated by a distance r attract each other according through gravity due to the force arising from a potential $V = -\frac{MmG}{r}$. It is convenient to take $m = 1$ equal to a "test mass," and to consider for the moment the potential due to a distribution of particles in the mass M. When we consider the action of the mass M on a second mass m, we replace the mass M by mM in the potential function.

Proposition 1 (Newton). *i. For a uniform sphere of mass M and radius a that is centered at the origin, the gravitational potential outside the sphere, namely at a distance $|\vec{x}| = r \geqslant a$ from the origin, is $V(\vec{x}) = -\frac{MG}{r}$. Here G is the gravitational constant. This is exactly the same as the potential of a point mass M located at the origin.*

ii. Consider a uniform shell of constant mass density ρ, and with radii $a_1 < a_2$ centered at the origin. Let $\vec{x}$ be a point enclosed by the shell, namely $0 \leqslant |\vec{x}| \leqslant a_1$. Then the potential $V(\vec{x}) = -2\pi\,(a_2^2 - a_1^2)\rho G$ is independent of $\vec{x}$ and yields no force.

iii. For a uniform sphere of mass M and radius a centered at the origin, its gravitational potential $V(r)$ at distance r from the center is

$$V(r) = \begin{cases} -\frac{3MG}{2a} + \frac{MG}{2a^3} r^2 & \text{for } r \leqslant a \\ -\frac{MG}{r} & \text{for } r \geqslant a \end{cases} . \tag{1}$$

Remark 2. The radial force corresponding to the potential (1) for a uniform sphere is

$$F_r = -\frac{dV(r)}{dr} = \begin{cases} -\frac{MG}{a^3}r & \text{for } r \leqslant a \\ -\frac{MG}{r^2} & \text{for } r \geqslant a \end{cases} . \tag{2}$$

However we will use this result for the rest of the course (ignoring the possibility of penetration into the interior of spherical masses) and consider the motion of spherical bodies point particles.

Proof. We present two proofs. The first method is computational; the second method is more conceptual. In the first proof we use spherical polar coordinates; if you have forgotten how they work, we review their use in §A1.

(i.) The mass density ρ of a uniform sphere of radius a (hence volume $\frac{4\pi a^3}{3}$) is defined by the relation "volume times mass density equals mass." Thus

$$M = \frac{4\pi a^3}{3}\rho . \tag{3}$$

By symmetry, it is sufficient to calculate the gravitational potential at a point $\vec{x}$ along the vertical (namely x_3) axis at a distance r from the center of the sphere. The potential at $\vec{x}$ due to an increment of volume $d\mathcal{V}$ that leads to an inverse square gravitational force is $-\frac{\rho G}{d}d\mathcal{V}$. Here d denotes the distance from $d\mathcal{V}$ to $\vec{x}$. Also let b denote the distance of $d\mathcal{V}$ from

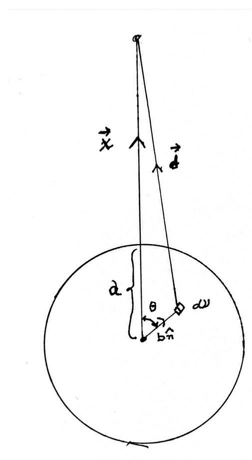

Figure 1: The potential $V(\vec{x})$ is the integral of contributions from volume elements $d\mathcal{V}$.

the origin, see Figure 1. Then the total potential at the $\vec{x}$ outside the sphere is the integral of the increment over the sphere, namely

$$\begin{aligned} V(\vec{x}) &= -\rho G \int_{\text{Sphere}} \frac{1}{d}\, d\mathcal{V} = -\rho G \int_0^{2\pi} d\varphi \int_0^a b^2 db \int_0^\pi \sin\theta d\theta\, \frac{1}{d} \\ &= -2\pi\rho G \int_0^a b^2 db \int_0^\pi \sin\theta d\theta\, \frac{1}{d}\,. \end{aligned} \tag{4}$$

We take b to be the radial coordinate and θ the angle between $d\mathcal{V}$ and the vertical axis. Thus we use $d\mathcal{V} = b^2 \sin\theta\, db\, d\theta\, d\varphi$. Since d does not depend on φ, that integral can be carried out and gives a factor 2π. To proceed, we need to know the dependence of d on b and θ, namely the function $d(b,\theta)$. We claim that

$$d(b,\theta) = \left(r^2 + b^2 - 2rb\cos\theta\right)^{\frac{1}{2}}\ , \tag{5}$$

where one takes the positive square root. This follows from the law of cosines. It can also be checked from squaring a vector identity. Let $\widehat{e}_3$ denote a unit vector in the $\vec{x}$ direction, so $\vec{x} = r\widehat{e}_3$. Also let $\widehat{n}$ denote a unit vector pointing from the origin to $d\mathcal{V}$. Finally let $\vec{d}$ denote the vector from $d\mathcal{V}$ to $\vec{x}$. Then $\vec{d} = \vec{x} - b\widehat{n}$. Squaring this yields (5), which put into (4) gives

$$\begin{aligned} V(\vec{x}) &= -2\pi\rho G \int_0^a b^2 db \int_0^\pi \frac{\sin\theta\, d\theta}{(r^2+b^2-2rb\,\cos\theta)^{\frac{1}{2}}} = -2\pi\rho G \int_0^a b^2 db \int_{-1}^1 \frac{ds}{(r^2+b^2-2rb\,s)^{\frac{1}{2}}} \\ &= \frac{2\pi\rho G}{r} \int_0^a b\,db\,(r^2+b^2-2rb\,s)^{\frac{1}{2}}\Big|_{s=-1}^{s=1} = \frac{2\pi\rho G}{r}\int_0^a b\,db\,((r-b)-(r+b)) \\ &= -\frac{4\pi\rho}{r} G \int_0^a b^2 db = -\left(\frac{4\pi a^3}{3}\rho\right)\frac{G}{r} = -\frac{MG}{r}\ , \end{aligned} \tag{6}$$

as claimed.

(ii.) The calculation for the potential at a point surrounded by the spherical shell is the same as the previous calculation, with two exceptions: 1.) The b integration $\int_{a_1}^{a_2} db$ is from a_1 to a_2. 2.) The positive square root that occurs at $s = 1$ is $(r^2+b^2-2rb\,s)^{\frac{1}{2}}\Big|^{s=1} = b - r$ and not $r - b$. This is the case because $r \leqslant a_1 \leqslant b \leqslant a_2$. Thus the calculation in (6) is replaced (starting in the second line) by

$$\begin{aligned} V(\vec{x}) &= \frac{2\pi\rho G}{r}\int_{a_1}^{a_2} b\,db\,(r^2+b^2-2rb\,s)^{\frac{1}{2}}\Big|_{s=-1}^{s=1} = \frac{2\pi\rho G}{r}\int_{a_1}^{a_2} b\,db\,((b-r)-(b+r)) \\ &= -4\pi\rho G \int_{a_1}^{a_2} b\,db = -2\pi\left(a_2^2 - a_1^2\right)\rho G\,. \end{aligned} \tag{7}$$

This is a constant independent of $\vec{x}$, so it has gradient zero, and it yields no force.

(iii.) The potential $V(r) = V_1(r) + V_2(r)$ for $r < a$ comes from adding the potential for a sphere of radius r given by part (i), namely

$$V_1(r) = -MG\frac{r^2}{a^3} \ ,$$

to the potential

$$V_2(r) = -MG\left(\frac{3}{2a} - \frac{3}{2}\frac{r^2}{a^3}\right) \ ,$$

for a shell with radii $r \leqslant a$ as given by part (ii). For $r > a$, $V(r)$ is given by part (i).

Second Proof: This second method might be regarded as more conceptual, as it is based on symmetry. It also is related to Gauss' law in electrostatics, and it relies on properties of the Laplacian, Green's theorem, and Poisson's equation. So if you have not studied these, please skip this second proof.[1]

(i.) Since the sphere has a uniform mass distribution, the potential at the point $\vec{x}$ does not change if we consider the sphere after a rotation about any axis through the center. Thus outside the sphere, the potential can only depend on the distance r from the origin, namely $V = V(r)$. Thus the gravitational force $\vec{F} = -\nabla V = -\frac{dV(r)}{dr}\widehat{n}$, where $\widehat{n}$ is a unit vector in the direction $\vec{x}$, has magnitude $\left|\vec{F}(r)\right|$ that is constant on a sphere S_r of radius $r \geqslant a$. Let $d\vec{A}$ denote the unit normal to a surface element for integration over a sphere. Hence by the representation of $V(\vec{x})$ as a solution to Poisson's equation $\Delta V = \rho$ for the mass density ρ, and using Green's theorem for the ball B_r bounded by S_r,

$$\begin{aligned} r^2\frac{dV(r)}{dr} &= \frac{1}{4\pi}\int_{S_r} d\vec{A}\cdot\nabla_x V(\vec{x}) = \frac{1}{4\pi}\int_{B_r} d^3x\,\Delta_x V(\vec{x}) = -\int_{B_r} d^3x\,\Delta_x\int\frac{\sigma(\vec{y})}{4\pi|\vec{x}-\vec{y}|}d^3y \\ &= \int_{B_r} d^3x\,\sigma(\vec{x}) = MG \ . \end{aligned} \tag{8}$$

Note that in our case $\sigma(\vec{x}) = \rho(\vec{x})G$, where $\rho(\vec{x})$ is the mass density of the uniform sphere of radius a. This is a constant ρ for $r \leqslant a$ and 0 for $r > a$. Thus $\int_{B_r} d^3x\,\sigma(\vec{x}) = MG$, and the gravitational force has magnitude $-\frac{dV(r)}{dr} = -\frac{MG}{r^2}$. The corresponding potential is $V(r) = -\frac{MG}{r}$, which agrees with (6). □

[1] In particular we use the fact that the Laplace operator in three dimensions is $\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}$ and has the property that $\Delta\left(\frac{1}{r}\right) = -4\pi\delta(\vec{x})$, where δ denotes the Dirac delta function. This "function" vanishes for $\vec{x} \neq 0$ and $\int\delta(\vec{x})d^3x = 1$. Here $d^3x = dx_1dx_2dx_3$. As a consequence, a spherically-symmetric solution to Poisson's equation $\Delta V(\vec{x}) = 4\pi\sigma(\vec{x})$ is $V(\vec{x}) = -\int\frac{\sigma(\vec{y})}{|\vec{x}-\vec{y}|}d^3y + C$. It is natural in this problem to put the 4π into the definition of $\sigma(\vec{x})$. Here C is a constant. If $\sigma(\vec{x})$ vanishes for large $|\vec{x}|$ as well as $V(\vec{x}) \to 0$, then $C = 0$, and we consider this case.

1.2 Point Particles

Consider a particle with mass m_1 at $\vec{x}_1$ and a second particle of mass m_2 at $\vec{x}_2$ interacting with the potential

$$V(\vec{x}_1, \vec{x}_2) = -\frac{k}{|\vec{x}_1 - \vec{x}_2|} \ . \tag{9}$$

If $k > 0$, the potential is attractive, while if $k < 0$, it is repulsive. This potential describes both gravitation and electrostatics. In case $k = Gm_1 m_2$, this describes Newtonian gravity with $G > 0$ the gravitational constant. In case $k = -q_1 q_2$, where q_j are electric charges, this describes the electrostatic Coulomb potential (opposite charges attract). In this lecture, we study this "two-body problem," although the generalization to N particles is straightforward. Here we study the attractive case $k > 0$.

This potential gives rise to the forces under Newton's law

$$\vec{F}_j(t) = m_j \vec{a}_j(t) = m_j \ddot{\vec{x}}_j(t) \ .$$

For each particle at $\vec{x}_j(t)$, the value of $\vec{x}_j$ at the instantaneous time t. And the relation between force and potential is

$$\vec{F}_j = -\nabla_{\vec{x}_j} V(\vec{x}_1, \vec{x}_2) \ .$$

Here ∇ denotes the gradient and $\nabla_{\vec{x}}$ the gradient with respect to the variable $\vec{x}$. At time t one has $\vec{F}_j(t) = -\nabla_{\vec{x}_j(t)} V(\vec{x}_1(t), \vec{x}_2(t))$.

One sees that the forces on the two particles are equal and opposite,

$$\vec{F}_1 = -\vec{F}_2 \ . \quad \text{or} \quad \vec{F}_1 + \vec{F}_2 = 0 \ . \tag{10}$$

In fact as the potential only depends on the distance $|\vec{x}_1 - \vec{x}_2|$ between the two particles,

$$\vec{F}_1 = -\nabla_{\vec{x}_1} V(\vec{x}_1, \vec{x}_2) = \nabla_{\vec{x}_2} V(\vec{x}_1, \vec{x}_2) = -F_2 \ .$$

2 The Two-Body Problem

We begin by showing that the two-body gravitational problem can be reduced to the study of a one-body problem. We begin by analyzing the center of mass coordinate $\vec{R}$. This is the average position, weighted by the mass. Equal and opposite forces on the two particles means that the *center of mass* moves freely.

The center of mass $\vec{R}$ is defined by the weighted average of the two positions $\vec{x}_1, \vec{x}_2$ in Cartesian coordinates,

$$\vec{R} = \frac{m_1}{m_1 + m_2}\, \vec{x}_1 + \frac{m_2}{m_1 + m_2}\, \vec{x}_2 \ .$$

The equation of motion for the center of mass $\vec{R}(t)$ due to the potential (9)–(10) is

$$\ddot{\vec{R}}(t) = \frac{m_1}{m_1+m_2}\,\ddot{\vec{x}}_1(t) + \frac{m_2}{m_1+m_2}\,\ddot{\vec{x}}_2(t) = \frac{1}{m_1+m_2}\left(\vec{F}_1 + \vec{F}_2\right) = 0\ .$$

As a consequence, the solution for the center of mass motion is

$$\vec{R}(t) = \vec{a} + \vec{b}\,t\ , \tag{11}$$

where the constant $\vec{a}$ is the position at time $t = 0$, and the constant $\vec{b}$ is the corresponding initial velocity at time $t = 0$. We can essentially ignore the motion of the center of mass and concentrate on understanding the motion of the two particles, relative to the center of mass.

2.1 Relative Coordinates

Introduce the relative coordinates $\vec{r}_j$, measured with respect to the position of the center of mass $\vec{R}$, as

$$\vec{r}_j = \vec{x}_j - \vec{R}\ , \quad \text{for } j = 1, 2\ . \tag{12}$$

We will solve for $\vec{r}_j$ as a function of time, which we denote by $\vec{r}_j(t)$. One can then recover the original coordinate $\vec{x}_j(t)$ as a function of time as

$$\vec{x}_j(t) = \vec{r}_j(t) + \vec{R}(t) = \vec{r}_j(t) + \vec{a} + \vec{b}t\ . \tag{13}$$

Notice that the definition of the center of mass ensures that the two coordinates $\vec{r}_j(t)$'s are related by the constraint

$$m_1\vec{r}_1(t) + m_2\vec{r}_2(t) = 0\ . \tag{14}$$

In other words, one of the three position vectors $\vec{R}, \vec{r}_1, \vec{r}_2$ is redundant.

Remark 3. In the case of two particles (as opposed to the N-particle generalization) it is easier to eliminate this redundancy immediately and to replace the two $\vec{r}_j$'s by the single relative displacement of one particle with respect to the other. Let

$$\vec{r} = \vec{r}_2 - \vec{r}_1\ , \quad \text{and denote} \quad r = |\vec{r}|\ , \quad \vec{n} = \frac{\vec{r}}{r}\ . \tag{15}$$

Denote the relative displacement $\vec{r}$ as a function of time by $\vec{r}(t)$. Then once we find $\vec{r}(t)$, we know both

$$\vec{r}_1(t) = -\left(\frac{m_2}{m_1+m_2}\right)\vec{r}(t)\ , \quad \text{and} \quad \vec{r}_2(t) = \left(\frac{m_1}{m_1+m_2}\right)\vec{r}(t)\ . \tag{16}$$

In other words, the *one-body problem* of determining the motion of $\vec{r}(t)$ is equivalent to solving the *two-body central force problem* of finding $\vec{r}_1(t)$ and $\vec{r}_2(t)$, or of finding $\vec{x}_1(t)$ and $\vec{x}_2(t)$.

2.2 Equations of Motion for the One-Body Problem

We will discover that the one body motion given by a central force lies in a plane, which is determined by two vectors: the position $\vec{r}$ and the velocity $\dot{\vec{r}}$. So before we analyze the motion, it is convenient to review briefly plane polar coordinates.

Plane Polar Coordinates We are interested in the relationship between Cartesian coordinates and plane polar coordinates. Cartesian coordinates in the plane refer to the standard coordinates: horizontal x_1 and vertical x_2. We denote the standard, fixed, orthonormal basis vectors by $\vec{e}_1$ and $\vec{e}_2$ oriented in the x_1 and x_2 directions respectively. Sometimes one called the coordinates x and y directions. A vector $\vec{x}$ in the plane with coordinates x_1, x_2 is written as

$$\vec{x} = x_1\,\vec{e}_1 + x_2\,\vec{e}_2\ . \tag{17}$$

Plane polar coordinates express $\vec{x}$ in terms of a different orthonormal basis that we denote

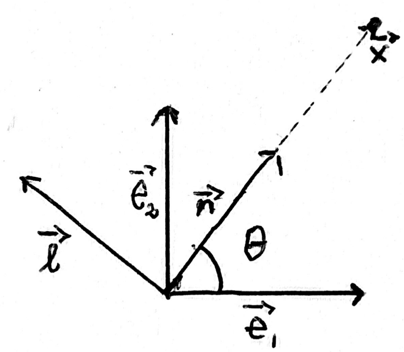

Figure 2: The Cartesian basis $\{\vec{e}_1, \vec{e}_2\}$ and the polar basis $\{\vec{n}, \vec{\ell}\}$ for a given vector $\vec{x}$.

by $\vec{n}$ and $\vec{\ell}$. Here $\vec{n}$ points from the origin toward $\vec{x}$, so both $\vec{n}$ and $\vec{\ell}$ depend on the vector $\vec{x}$. In other words, using the standard convention that the coordinate r denotes the length $\|\vec{x}\|$ of $\vec{x}$ by $\|\vec{x}\| = \sqrt{x_1^2 + x_2^2}$, one has

$$\vec{x} = r\vec{n}\ . \tag{18}$$

Also the coordinate θ denotes the angle between $\vec{x}$ and the x_1 axis, namely $\vec{n}\cdot\vec{e}_1 = \cos\theta$, and we choose the standard orientation for the sign of θ, namely increasing θ rotates $\vec{n}$ counter clockwise. Then

$$\vec{n} = \cos\theta\,\vec{e}_1 + \sin\theta\,\vec{e}_2\,, \quad \vec{\ell} = -\sin\theta\,\vec{e}_1 + \cos\theta\,\vec{e}_2\,. \tag{19}$$

As $\vec{n}$ and $\vec{\ell}$ are given functions of polar angle θ, and as the Cartesian basis is fixed, it is easy to track how $\vec{n}$ and $\vec{\ell}$ change with time. Thus[2]

$$\frac{d\vec{n}}{d\theta} = \vec{\ell}\,, \quad \frac{d\vec{\ell}}{d\theta} = -\vec{n}\,. \tag{20}$$

Using the chain rule $\frac{d\vec{n}(\theta(t))}{dt} = \frac{d\vec{n}(\theta(t))}{d\theta(t)}\dot{\theta}(t)$, and abbreviating $\frac{d\vec{n}(\theta(t))}{dt}$ as $\dot{\vec{n}}$, we find that

$$\boxed{\dot{\vec{n}} = \dot{\theta}\,\vec{\ell}\,, \quad \dot{\vec{\ell}} = -\dot{\theta}\,\vec{n}}\,. \tag{21}$$

Derivation of the One-Body Equation We claim that the coordinates $\vec{r}(t)$ satisfy the equations of motion

$$\dot{\vec{p}} = \mu\ddot{\vec{r}} = \vec{F}\,, \quad \text{where} \quad \vec{F} = -\nabla_{\vec{r}}V(r) = -\frac{k}{r^3}\,\vec{r}\,. \tag{22}$$

Here the quantity μ is the *reduced mass* defined as

$$\boxed{\mu = \tfrac{m_1 m_2}{m_1+m_2}\,, \quad \text{or} \quad \tfrac{1}{\mu} = \tfrac{1}{m_1} + \tfrac{1}{m_2}}\,. \tag{23}$$

In the limit that $m_1 \to \infty$, one has $\mu = m_2$. In other words, one can ignore the motion of an infinitely heavy particle.[3]

In order to derive this result, note that

$$m_1\ddot{\vec{x}}_1 = -\nabla_{\vec{x}_1}V(r) = \frac{k}{r^2}\,\vec{n}\,, \quad \text{and} \quad m_2\ddot{\vec{x}}_2 = -\nabla_{\vec{x}_2}V(r) = -\frac{k}{r^2}\,\vec{n}\,. \tag{24}$$

Here we use the identity $r = \left(\sum_{j=1}^{3}\left(x_{2j} - x_{1j}\right)^2\right)^{1/2}$ giving

$$\frac{\partial r}{\partial x_{1j}} = \frac{x_{1j} - x_{2j}}{r} = -\vec{n}_j\,, \quad \text{and} \quad \frac{\partial r}{\partial x_{2j}} = \frac{x_{2j} - x_{1j}}{r} = \vec{n}_j\,. \tag{25}$$

Define the difference coordinate $\vec{r} = \vec{x}_2 - \vec{x}_1$. Using (24) and (23) gives (22), namely

$$\ddot{\vec{r}} = \ddot{\vec{x}}_2 - \ddot{\vec{x}}_1 = -\left(\frac{1}{m_2} + \frac{1}{m_1}\right)\frac{k}{r^2}\,\vec{n} = -\frac{k}{\mu r^2}\,\vec{n}\,. \tag{26}$$

[2]Note that the vector $\vec{n}(\theta)$ is a unit vector, so $\frac{d\vec{n}(\theta)}{d\theta}$ must be orthogonal to $\vec{n}(\theta)$. However, the fact that the derivative is also a unit vector depends on the particular parameterization (19). The same holds for $\vec{\ell}(\theta)$.

[3]In this limit the center of mass becomes the position of the heavy particle, $\vec{R} = \vec{x}_1$. And unless this particle remains at rest, its kinetic energy is infinite. So we assume an infinitely heavy particle remains at rest. In this case, the relative coordinate $\vec{r}(t)$ is just equal to $\vec{r}_2(t)$.

3 Three Conservation Laws and $\frac{1}{r^2}$ Forces

A quantity is said to be *conserved*, if it remains constant under time evolution. We claim that three basic quantities are conserved under the evolution in the two-body Kepler problem. They are the energy, the angular momentum vector, and the Runge-Lenz vector. Energy is conserved for all forces that are the gradient of a potential, and angular momentum is conserved for all central force laws. **But conservation of the Runge-Lenz vector is special for an inverse-square central force.**

3.1 Energy

One says a force $\vec{F} = -\nabla V$ is *conservative*, as the energy

$$E = T + V = \frac{\vec{p}^2}{2\mu} + V \tag{27}$$

is conserved. Using Newton's law, $\mu\ddot{\vec{r}} = \vec{F}$, one sees that energy is conserved whenever the force comes from a potential:

$$\frac{dE}{dt} = \mu\ddot{\vec{r}} \cdot \dot{\vec{r}} + \nabla V \cdot \dot{\vec{r}} = \left(\mu\ddot{\vec{r}} - \vec{F}\right) \cdot \dot{\vec{r}} = 0 \ .$$

3.2 Angular Momentum

The angular momentum about the origin is defined by $\vec{L} = \vec{r} \times \vec{p}$, where $\vec{p} = \mu\dot{\vec{r}}$. A *central force potential* V is a potential whose gradient is parallel to $\vec{r}$. We claim that $\vec{L}$ is conserved in any problem with a central force potential. In order to see this, look at

$$\frac{d\vec{L}}{dt} = \frac{d}{dt}(\vec{r} \times \vec{p}) = \dot{\vec{r}} \times \vec{p} + \vec{r} \times \vec{F} = 0 + 0 = 0 \ . \tag{28}$$

The first term vanishes because $\dot{\vec{r}} = \mu^{-1}\vec{p}$ is in the same direction as $\vec{p}$. The second term vanishes because $\vec{F}$ is in the same direction as $\vec{r}$. In the case of the potential $V = -\frac{k}{r}$, the force is $\vec{F} = -\frac{k}{r^2}\vec{n} = -\frac{k}{r^3}\vec{r}$.

Since $\vec{L}$ is constant, it determines a fixed plane orthogonal to $\vec{L}$. We can describe the planar motion in terms of planar polar coordinates. Then $\vec{p} = \mu\dot{r}\vec{n} + \mu r\dot{\theta}\,\vec{\ell}$, and consequently

$$\vec{L} = \mu r^2\dot{\theta}\ \vec{n} \times \vec{\ell}\ , \tag{29}$$

The magnitude of $\vec{L}$ is

$$L = \mu r^2|\dot{\theta}|\ . \tag{30}$$

Conservation of $\vec{L}$ means that the motion lies in the plane perpendicular to $\vec{L}$.

3.3 The Runge-Lenz Vector

Define the Runge-Lenz vector $\vec{\epsilon}$ in terms of the momentum $\vec{p} = \mu\dot{\vec{r}}$ and $\vec{L}$ by

$$\boxed{\vec{\epsilon} = \tfrac{\vec{p}\times\vec{L}}{\mu k} - \vec{n}}\,. \tag{31}$$

Interpretation: Note that both terms in $\vec{\epsilon}$ lie in the plane of motion, so $\vec{\epsilon}$ does as well. In the end, we will interpret $\vec{\epsilon}$ as a vector describing the eccentricity of the motion. It points along the semi-major axis of motion, and its magnitude is the classical eccentricity of the orbit. For this reason we have defined $\vec{\epsilon}$ to be a dimensionless quantity.

We now show that $\vec{\epsilon}$ is conserved. The only time-dependent quantities in (31) are $\vec{p}$ and $\vec{n}$. Using $\dot{\vec{p}} = \vec{F}$,

$$\dot{\vec{\epsilon}} = \frac{\dot{\vec{p}}\times\vec{L}}{\mu k} - \dot{\vec{n}} = \frac{\vec{F}\times\vec{L}}{\mu k} - \dot{\vec{n}}\,. \tag{32}$$

The vector identity

$$\vec{A}\times(\vec{B}\times\vec{C}) = (\vec{A}\cdot\vec{C})\vec{B} - (\vec{A}\cdot\vec{B})\vec{C}$$

shows $\vec{n}\times(\vec{n}\times\vec{\ell}) = -\vec{\ell}$. Recall (29), and (21), so

$$\vec{F} = -\frac{k}{r^2}\,\vec{n}\,,\quad \vec{L} = (\mu r^2\dot{\theta})\,\vec{n}\times\vec{\ell}\,,\quad \text{and}\quad \dot{\vec{n}} = \dot{\theta}\,\vec{\ell}\,.$$

Hence $\vec{F}\times\vec{L} = \mu k\ \dot{\vec{n}}$. We conclude from (32) that $\dot{\vec{\epsilon}} = 0$. Hence $\vec{\epsilon}$ is conserved as claimed.

It is natural to ask whether $\vec{\epsilon}$ is independent of the conservation laws for E and $\vec{L}$ that we already discovered. Clearly ϵ is perpendicular to $\vec{L}$, so it must lie in the plane of motion (and hence is not completely independent). In what direction does $\vec{\epsilon}$ point? The simplest thing is to say that we orient our coordinates so that $\vec{\epsilon}$ points in a standard direction, say the x_1 axis. In that case, the angle θ between $\vec{e}p$ and $\vec{r}$ is the same as the angle θ in plane polar coordinates. So we will discuss that case. But in general the vector ϵ can point in any direction with angle θ_0 to the real axis.

Let us see that the length ϵ of the vector $\vec{\epsilon}$ is a function of E and $\vec{L}$ is

$$\boxed{\epsilon = \left(1 + \frac{2EL^2}{\mu k^2}\right)^{1/2}}\,. \tag{33}$$

Note that $(\vec{p}\times\vec{L})^2 = p^2\,L^2$, as $\vec{p}\perp\vec{L}$. So using $\left(\vec{A}\times\vec{B}\right)\cdot\vec{C} = \vec{A}\cdot\left(\vec{B}\times\vec{C}\right)$, see (79),

$$\frac{2}{\mu k}\vec{n}\cdot\left(\vec{p}\times\vec{L}\right) = \frac{2}{\mu k r}\vec{r}\cdot\left(\vec{p}\times\vec{L}\right) = \frac{2}{\mu k r}\,(\vec{r}\times\vec{p})\cdot\vec{L} = \frac{2L^2}{\mu k^2}\frac{k}{r}\,.$$

Hence

$$\vec{\epsilon}^{\,2} \;=\; 1+\frac{p^2L^2}{\mu^2k^2}-\frac{2}{\mu k}\vec{n}\cdot\left(\vec{p}\times\vec{L}\right)=1+\frac{2L^2}{\mu k^2}\left(\frac{p^2}{2\mu}-\frac{k}{r}\right)=1+\frac{2EL^2}{\mu k^2}\;. \tag{34}$$

4 Equation of the Orbit from the Conservation Laws

The equation of the orbit gives $\vec{r}$ as a function of the angle θ; it is an immediate consequence of the conservation laws for $\vec{\epsilon}$, $\vec{L}$ and for E. Note using (79) and $\vec{n}=\frac{\vec{r}}{r}$, that

$$\vec{r}\cdot\vec{\epsilon}=\frac{L^2}{\mu k}-r \tag{35}$$

Let θ denote the angle between $\vec{r}$ and $\vec{\epsilon}$. Thus $\vec{r}\cdot\vec{\epsilon}=r\epsilon\cos\theta$, yielding the formula for a conic section

$$\boxed{r\left(1+\epsilon\cos\theta\right)=\frac{L^2}{\mu k}}\;. \tag{36}$$

The quantity $L^2/\mu k$ on the right side of (36) is a constant. Here ϵ is the length of the Runge-Lenz vector $\vec{\epsilon}$, so it is natural to call the Runge-Lenz vector the *vector eccentricity*. In geometry the length ϵ is called the *eccentricity* of the orbit. If $\epsilon=0$ the motion is circular. For $0<\epsilon<1$ the motion is an ellipse with different semi-major and semi-minor axes. If $1<\epsilon$, the motion is on a hyperbola, and θ has a maximum value. The vector $\vec{\epsilon}$ points along the major axis of the ellipse.

In case the ellipse is oriented at a different angle to the x_1 axis, this just means that in terms of the polar angle θ the angle between $\vec{r}$ and $\vec{\epsilon}$ is $\theta-\theta_0$, so orbit would be

$$\boxed{r\left(1+\epsilon\cos(\theta-\theta_0)\right)=\frac{L^2}{\mu k}}\;. \tag{37}$$

For the rest of this section, let us assume that $\theta_0=0$. so that $\vec{\epsilon}$ points along the x_1 axis. In fact this is the direction of the major axis of the motion, namely the direction (for $\epsilon<1$) where the value of r takes on both its maximum and its minimum values.

Note the relation between the ellipse (36) for $\epsilon<1$ and the corresponding formula in Cartesian coordinates is worked out in §A3.1. The corresponding calculation for a parabola ($\epsilon=1$) is in §A3.2, and for a hyperbola ($\epsilon>1$) in §A3.3. In the case of elliptic orbits, we show in §A3.1 that the ellipse expressed in Cartesian coordinates as

$$\left(\frac{x_1}{a}\right)^2+\left(\frac{x_2}{b}\right)^2=1\;. \tag{38}$$

for $0 < b \leqslant a$ corresponds to the elliptical orbit

$$r(1+\epsilon\cos\theta) = a(1-\epsilon^2)\ , \tag{39}$$

with

$$\epsilon = \sqrt{1-\left(\frac{b}{a}\right)^2}\ . \tag{40}$$

In terms of the physical parameters,

$$a = -\frac{k}{2E}\ , \quad b = \frac{L}{\sqrt{-2\mu E}}\ . \tag{41}$$

While we have chosen to orient the coordinates so for $\theta_0 = 0$ the vector $\vec{\epsilon}$ points along the x_1 axis, it is nice to check from the definition that this is the case. We can verify that for $\epsilon \neq 0$, the vector $\vec{\epsilon}$ points along the $\vec{e}_1$ direction, i.e. along semi-major axis of motion. One sees this by evaluating $\vec{\epsilon}$ at two different points on the orbit. If we plug in the values of $\vec{p}, \vec{L}, \widehat{n}$ at $\theta = 0$, we conclude its direction is either $\pm\vec{e}_1$, for in this case, $\vec{p}\times\vec{L}$ points in the direction $\vec{e}_1$ while $-\vec{n}$ points in the opposite direction. In any case, ϵ has no component orthogonal to $\vec{e}_1$. Alternatively we can evaluate $\vec{\epsilon}$ for an elliptical orbit at the point on the orbit where (the Cartesian coordinate) $x_1 = 0$. This is the point where $\vec{r}$ has the largest value of x_2, and at that point $\vec{p}\times\vec{L}$ points along the semi-minor axis in the direction $\vec{e}_2 \perp \vec{e}_1$. So $\vec{p}\times\vec{L}$ has zero component in the $\vec{e}_1$ direction. At this point $\theta > \pi/2$ so $\vec{r}$ has a negative component in the $\vec{e}_1$ direction, so $\vec{\epsilon}$ has a positive component in the $\vec{e}_1$ direction.

One sees this by evaluating $\vec{\epsilon}$ when $\theta = 0$. In this case both $\vec{n}$ and $\vec{p}\times\vec{L}$ point in the same direction. And as $\theta = 0$, this direction is along the semi-major axis. A circular orbit has eccentricity $\epsilon = 0$. The energy of a circular orbit is the minimal energy for all orbits with given constants μ, k; a circular orbit has energy $E = -\frac{\mu k^2}{2L^2}$. Elliptical orbits have eccentricity in the interval $0 \leqslant \epsilon < 1$, with the corresponding negative energy in the interval $-\frac{\mu k^2}{2L^2} \leqslant E < 0$.

4.1 The Effective Potential

Using plane polar coordinates, $\dot{\vec{r}} = \dot{r}\,\vec{n} + r\dot{\theta}\,\vec{\ell}$. Thus

$$\dot{\vec{r}}^{\,2} = \dot{r}^2 + r^2\dot{\theta}^2 = \dot{r}^2 + \frac{L^2}{\mu^2 r^2}\ .$$

In the last equality we eliminate $\dot{\theta}$ in terms of r and the conserved quantity L by using (30). One then can write the energy E as

$$\begin{aligned} E &= T + V = \frac{1}{2}\mu\dot{\vec{r}}^{\,2} - \frac{k}{r} \\ &= \frac{1}{2}\mu\dot{r}^2 + \frac{L^2}{2\mu r^2} - \frac{k}{r} = \frac{1}{2}\mu\dot{r}^2 + V_{\text{eff}}(r)\ . \end{aligned} \tag{42}$$

Thus conservation of angular momentum can be used to cast the energy in the form of a radial kinetic energy plus an effective potential for radial motion

$$V_{\text{eff}}(r) = \frac{L^2}{2\mu r^2} - \frac{k}{r}\ . \tag{43}$$

The effective potential has the form in Figure 3: In Figure 3 the wavy line represents the

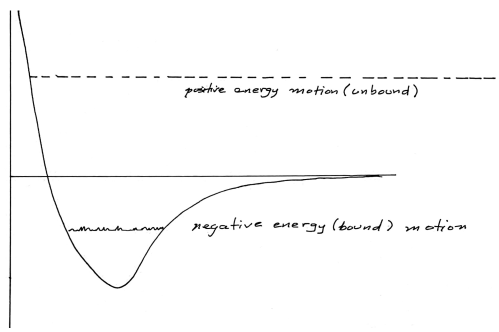

Figure 3: Effective potential $V_{\text{eff}}(r) = \frac{L^2}{2\mu r^2} - \frac{k}{r}$, in the attractive case $k > 0$.

energy of a bound state solution. In this case the value of r lies between a minimum value $r_{\text{min}} = \frac{L^2}{\mu k(1+\epsilon)}$ and maximum value $r_{\text{max}} = \frac{L^2}{\mu k(1-\epsilon)}$ where $\dot{r}$ changes sign, On the other hand, the dotted line denotes the energy of an unbound orbit.

5 Kepler's Laws

Kepler observed planetary motion, and formulated several things known as his *laws.*

1. *The orbits are ellipses.* We have already verified this.

2. *Equal areas are swept out in equal time.*

Consider the area swept out by the vector from one focus to the orbiting particle. The increment in area $dA = \frac{1}{2}r^2 d\theta$ (with the origin of coordinates at the focus) leads to the constant change in area under time evolution, as the magnitude of the angular momentum is $L = \mu r^2 \dot{\theta}$. Then

$$\frac{dA}{dt} = \frac{1}{2} r^2 \dot{\theta} = \frac{L}{2\mu} \,, \tag{44}$$

3. *The periods τ of the planets are proportional to $a^{3/2}$, where a is the semi-major axis.*

If τ is the period of the orbit, then the total area of the orbit is

$$A = \frac{L\tau}{2\mu} \,, \qquad \text{or } \tau = \frac{2\mu A}{L} \,. \tag{45}$$

The area of an ellipse is given by the length of its axes, and it equals $A = \pi ab$. Inserting the relations (84) and (88), we see that

$$\tau = \frac{2\mu A}{L} = \frac{2\mu\pi ab}{\sqrt{-2\mu E}\; b} = 2\pi \sqrt{\frac{\mu}{k}} a^{3/2} \,. \tag{46}$$

In the Kepler problem,

$$k = G m_1 m_2 \,, \quad \text{while} \quad \mu = \frac{m_1 m_2}{m_1 + m_2} \,. \tag{47}$$

Take m_1 as the mass of the sun and m_2 the mass of the planet, so

$$\sqrt{\frac{\mu}{k}} = \frac{1}{\sqrt{G m_1}} \left(1 + \frac{m_2}{m_1}\right)^{-1/2} \,, \tag{48}$$

and

$$\tau = \frac{2\pi}{\sqrt{G m_1}} \left(1 + \frac{m_2}{m_1}\right)^{-1/2} a^{3/2} \,. \tag{49}$$

For $m_2 \ll m_1$, the coefficient $(1 + \frac{m_2}{m_1})^{-1/2} \sim 1 - \frac{m_2}{2m_1}$ is approximately equal to 1. To the extent this coefficient is constant, the periods τ of the various planets (with different m_2's) approximately obey Kepler's third law as a function of a.

6 Solution for the Orbit Using Newton's Equation

When we considered central force motion, we wrote a vector in the plane in the form

$$\vec{r} = r\vec{n} \,, \tag{50}$$

where we used the orthonormal basis $\vec{n}, \vec{\ell}$. Then we resolved the velocity into its radial and angular components using $\dot{\vec{n}} = \dot{\theta}\,\vec{\ell}$, so

$$\dot{\vec{r}} = \dot{r}\,\vec{n} + r\dot{\theta}\,\vec{n} \,. \tag{51}$$

Likewise using $\dot{\vec{\ell}} = -\dot{\theta}\,\vec{n}$ we can resolve the acceleration as

$$\ddot{\vec{r}} = (\ddot{r} - r\dot{\theta}^2)\vec{n} + (2\dot{r}\dot{\theta} + r\ddot{\theta})\vec{\ell} = (\ddot{r} - r\dot{\theta}^2)\vec{n} + \frac{1}{r}\left(\frac{d}{dt}\left(r^2\dot{\theta}\right)\right)\vec{\ell} \,. \tag{52}$$

The term $\ddot{r}$ is called the "radial acceleration;" the term $r\dot{\theta}^2$ is called the "centripetal acceleration;" and the final term is called the "torque." Multiplying by mass μ, we obtain the equation of a planar orbit by using Newton's equation $\vec{F} = \mu\vec{a}$. In the case of the Kepler problem, $\vec{F} = -\frac{k}{r^2}\vec{n}$, we get

$$\mu(\ddot{r} - r\dot{\theta}^2)\vec{n} + \frac{d}{dt}\,(\mu)\,r^2\dot{\theta}\vec{\ell} = -\frac{k}{r^2}\vec{n} \,. \tag{53}$$

The radial component of (53) is the equation is

$$\mu(\ddot{r} - r\dot{\theta}^2) = -\frac{k}{r^2} \,, \tag{54}$$

while the angular equation is

$$\frac{dL}{dt} = 0 \,, \quad \text{where} \quad L = \mu r^2\dot{\theta} \,. \tag{55}$$

The angular equation gives conservation of angular momentum. We can use that to eliminate $\dot{\theta}$ from the radial equation, replacing it with

$$\dot{\theta} = \frac{L}{\mu r^2} \,, \quad \text{so} \quad \mu\ddot{r} - \frac{L^2}{\mu r^3} = -\frac{k}{r^2} \,. \tag{56}$$

One can solve the radial equation for $r(\theta)$ by making the substitution $r = \frac{1}{u}$. Here the variable u is a function of θ, and θ is a function of t. Let u' denote the derivative of u with respect to θ, namely $u' = \frac{du}{d\theta}$. Then

$$\dot{u} = u'\dot{\theta} = u'\frac{L}{\mu r^2} = \frac{L}{\mu}u^2u' \,. \tag{57}$$

Returning to the radial equation, and the variable $r = \frac{1}{u}$,

$$\dot{r} = -\frac{1}{u^2}\dot{u} = -\frac{L}{\mu}u' \ , \quad \text{so} \quad \ddot{r} = -\frac{L}{\mu}u''\dot{\theta} = -\frac{L^2}{\mu^2}u^2u'' \tag{58}$$

Putting this back into the radial equation (56), we obtain

$$-\frac{L^2}{\mu}u^2u'' - \frac{L^2}{\mu}u^3 = -ku^2 \ .$$

Multiply by $-\frac{\mu}{L^2u^2}$ to give

$$\boxed{u'' + u = \tfrac{\mu k}{L^2}} \ . \tag{59}$$

Equation (59) is an "inhomogeneous second order ordinary differential equation." One obtains the general solution for u by adding to the particular solution $u = \frac{\mu k}{L^2}$ the general solution to the homogeneous equation $u'' + u = 0$. The homogeneous equation has solutions given by constants ϵ, α to match $u(0)$ and $u'(0)$, namely $u = \epsilon\frac{\mu k}{L^2}\cos(\theta - \alpha)$. So

$$u = \frac{\mu k}{L^2}\left(1 + \epsilon\cos(\theta - \alpha)\right) = \frac{1}{r} \ . \tag{60}$$

With $\alpha = 0$ the major axis is horizontal, and we have recovered our previous solution given in (36).

7 Unbound Orbits

The hyperbola is defined as the locus of points such that the difference of the distances $r = |\vec{r}|$ and $r' = \left|\vec{r}'\right|$ to two fixed foci remain constant. The hyperbola defined in this way has two branches, depending on whether one considers $r' - r$ or $r - r'$. The relation between the discussion in polar coordinates and the representation of the hyperbola as a conic section in Cartesian coordinates is given in detail in §A3.3.

7.1 Attractive Potentials and Unbound Orbits

The analysis is similar to the bound case. The only difference is that now the energy $E > 0$ is positive, so the eccentricity $\epsilon = \left(1 + \frac{2EL^2}{\mu k^2}\right)^{\frac{1}{2}}$ satisfies $\epsilon > 1$. The orbit is described by

$$r\left(1 + \epsilon\cos\theta\right) = \frac{L^2}{\mu k} \ . \tag{61}$$

Thus as θ increases, there will be a maximum value, namely $\cos\theta_{\max} = -1/\epsilon$, at which $r \to \infty$. We claim that the vector $\vec{\epsilon}$ points from the focus to the closest point on the orbit, so in the diagram Figure 10 the vector $\vec{\epsilon}$ points to the right. The easiest way to see this is to compute $\vec{\epsilon}$ again at two values of θ. As in the bound case, at $\theta = 0$ one concludes that $\vec{\epsilon}$ points from the focus along the x_1 axis, either toward or away from the particle. Since $\frac{\pi}{4} < \theta_{\max}$, see (97), computing at any point for which $\frac{\pi}{4} < \theta < \theta_{\max}$ shows that both $\vec{p} \times \vec{L}$ and $-\vec{n}$ have positive components along the $\theta = 0$ axis.

7.2 Repulsive Potentials and Unbound Orbits

We take the potential in the repulsive one-body problem, reduced from the two-body problem for two charged particles having charges q_1, q_2 of the same sign. Then $k = -q_1 q_2$ (in CGS units, while in other units k includes some other constant). Thus the potential is

$$V(r) = -\frac{k}{r} = \frac{q_1 q_2}{r} > 0 ,$$

with the sign indicating a repulsive force. The effective potential is strictly positive and monotonically decreasing,

$$V_{\text{eff}}(r) = \frac{L^2}{2\mu r^2} + \frac{|k|}{r} , \tag{62}$$

so its derivative is strictly negative, indicating that the radial force is strictly repulsive.

As in the attractive case, there are three conserved quantities: the energy E, and two vectors, the angular momentum $\vec{L}$, and the Runge-Lenz vector $\vec{\epsilon}$. In terms of formulas, the conserved quantities are

$$E = \frac{p^2}{2\mu} + \frac{|k|}{r} , \qquad \vec{L} = \vec{r} \times \vec{p} , \qquad \vec{\epsilon} = \frac{\vec{p} \times \vec{L}}{\mu k} - \vec{n} = -\frac{\vec{p} \times \vec{L}}{\mu |k|} - \vec{n} . \tag{63}$$

In the attractive case, the vector $\vec{\epsilon}$ points from the focus along the semi-major axis in the direction $\vec{e}_1$. In the repulsive case, we claim that the vector $\vec{\epsilon}$ points in the opposite direction $-\vec{e}_1$. In fact, if we evaluate $\vec{\epsilon}$ at the point of the orbit when $\theta = 0$, both terms in the formula for $\vec{\epsilon}$ in (63) point in the direction $-\vec{e}_1$.

7.3 Equation of the Orbit

The derivation of the hyperbolic orbit for unbound motion is the same as the bound case. However, in our choice of polar coordinates, we use a different definition of the angle θ. In the bound case the motion crosses the x_1 axis at $\theta = 0$, and if its motion at that point increases x_2, then $\vec{L} = \vec{r} \times \vec{p}$ points in the positive x_3 direction. We also saw that in this case ϵ points in the direction $\vec{e}_1$. As previously, we place the origin of coordinates on the x_1 axis. However in the unbound case, the change of sign of k means that $\vec{\epsilon}$ points in the opposite direction $-\vec{e}_1$. So in the unbound case, we define the angle θ as the angle between the vector $\vec{r}$ and $\vec{e}_1$ which differs from the angle between $\vec{\epsilon}$ and $\vec{e}_1$ by π. The equation for the orbit is then

$$\vec{r} \cdot \vec{\epsilon} = -r\epsilon \cos\theta = \vec{r} \cdot \left(-\frac{\vec{p} \times \vec{L}}{\mu|k|} - \vec{n} \right) = -\frac{L^2}{\mu|k|} - r \; , \tag{64}$$

or

$$r\left(-1 + \epsilon\cos\theta\right) = \frac{L^2}{\mu|k|} \; , \tag{65}$$

which is the parametric form of the negative branch of a hyperbola. We illustrate the orbit. In this picture, the two foci are denoted F, F' and $\theta_{\max}$ is denoted by α. The force center is located at F.

As in the bound case, the eccentricity is given by squaring the vector $\vec{\epsilon}$, with the same result, namely

$$\epsilon^2 = 1 + \frac{2EL^2}{\mu k^2} \; . \tag{66}$$

A hyperbolic orbit along the negative branch is illustrated in Figure 4, along with the straight line asymptotes which are tangents to the orbit in the limit as it goes to $r = \infty$. The asymptotes to the orbit cross at the origin of the Cartesian coordinate system, which is placed midway between the two foci. Each asymptote lies at a perpendicular distance s from a second line parallel to the asymptote, and passing through the origin of coordinates at the force center (focus). The distance s is called the *impact parameter*. This measures how far the incoming particle would miss the target located at the focus, were it to travel on straight line motion, uninfluenced by the force, and parallel to the asymptote of the incoming trajectory. The equation for the the orbit (65) shows that $-1 + \epsilon\cos\theta_{\max} = 0$, or

$$\cos\theta_{\max} = \frac{1}{\epsilon} \; . \tag{67}$$

One can rotate the orbit in Figure 4 into the standard position to describe the scattering experiment, shown in Figure 5.

8 The Rutherford Scattering Experiment

To describe the experiment, one turns the picture around in order to describe a beam of particles coming toward the target. **Warning:** *In the consideration of Rutherford scattering,*

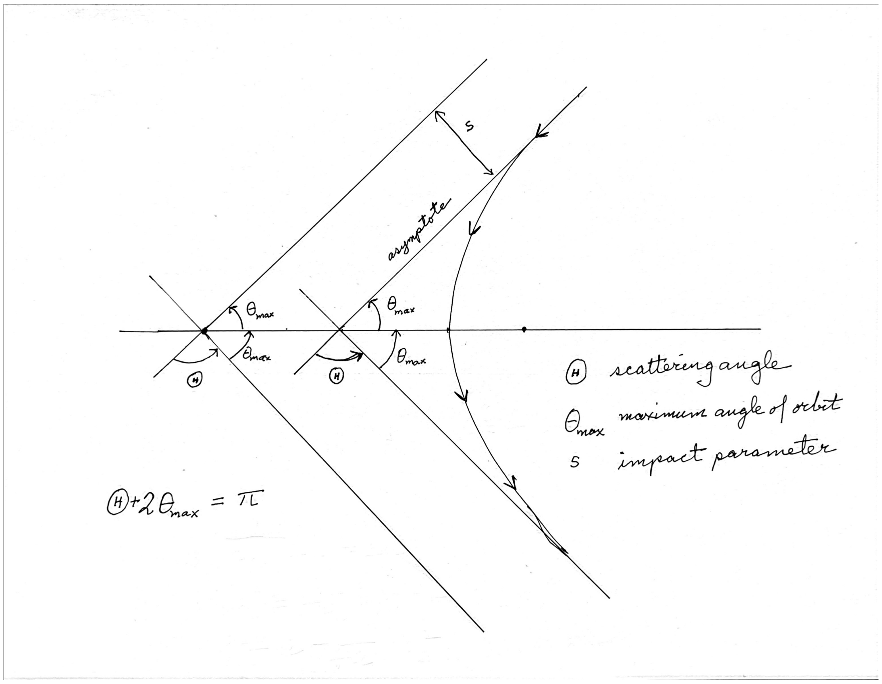

Figure 4: Repulsive Scattering Orbit

the z (i.e. the x_3) axis is rotated to be in the x (i.e. the x_1) direction, and the angle denoted here by Θ is the scattering angle. This is not the angle θ in the equation of the orbit, but these two angles are related as we see below.

This diagram shows that the analysis of the experiment is a purely geometric question. The incoming beam transverses the plane orthogonal to it. It is deflected by the target at the force center; this is denoted F, and it is the *center* of a large sphere around the target. The target at F is also the *focus* of the hyperbola describing the orbit.

8.1 The Differential Scattering Cross Section

The Rutherford experiment was to send a beam of "alpha particles," which are positively-charged helium nuclei, to scatter on a gold film. If the nuclear charge is spread out over a spherical atom, then the scattering would be like shooting a beam into a mush. That is not what Rutherford and his colleagues observed. They found the incoming positively charged beam was deflected as if the repulsive center of the gold film was concentrated in each atom in a tiny volume at the center–namely in the nucleus. The angle of deflection from the repulsive

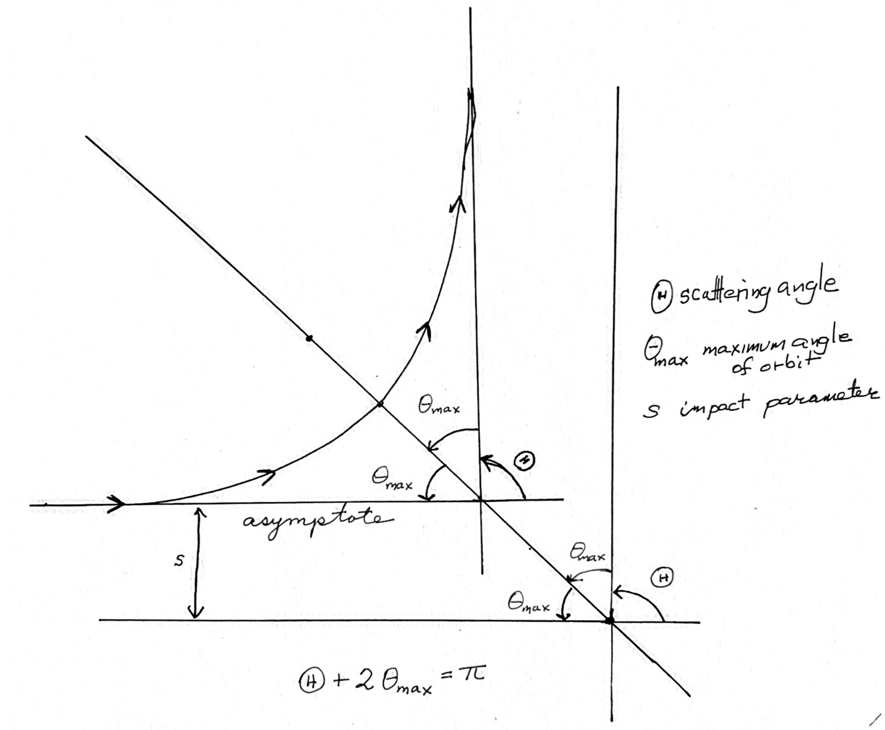

Figure 5: Rutherford Experiment Oriented Horizontally

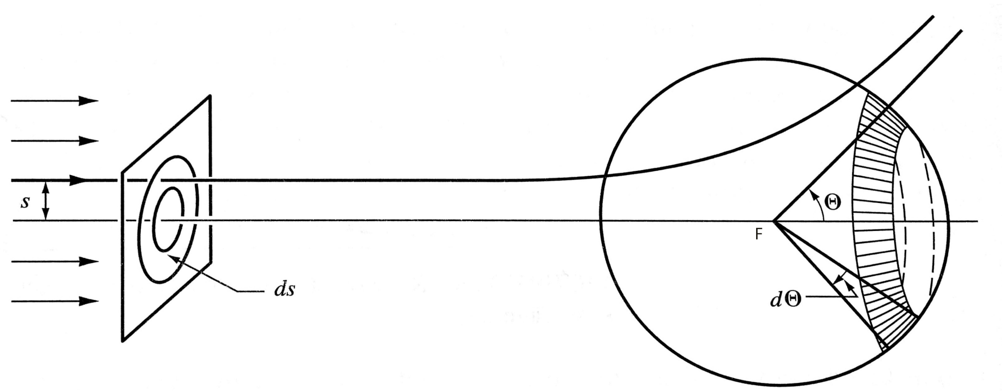

Figure 6: Idea of the Experiment (from p. 107 in Goldstein, Poole, and Safko).

$1/r^2$ electro-static force depends on the perpendicular distance of the incoming beam from a direct line to the repulsive force center. The gold atoms are very heavy compared with the helium projectiles, so they can be considered approximately fixed at the repulsive focus of the orbit.

The differential scattering cross section is the ratio of the area of an annulus in the incoming beam to the solid angle into which it scatters. It is traditional to denote this cross section as $\frac{d\sigma}{d\Omega}$. The most relevant parameter is the *scattering angle* Θ. This is the angle between the two asymptotes of the orbit and is illustrated in Figures 4–6.

We now show that for an incoming beam of particles, each of which has energy E, that the Rutherford cross section is the expression:

$$\boxed{\frac{d\sigma}{d\Omega} = \left(\frac{k}{4E}\right)^2 \frac{1}{\sin^4 \frac{\Theta}{2}}}\,. \tag{68}$$

This result is amazing as it is completely classical, while the phenomenon being observed is ultimately described by quantum mechanics. It is a lucky scientific accident that the answer from quantum theory is almost the same as the classical answer. Othewise Rutherford may not have discovered the nucleus!

To derive this formula, we can use polar coordinates

$$\frac{d\sigma}{d\Omega} = \left|\frac{s ds d\varphi}{d\Omega}\right| = \frac{1}{\sin\Theta}\, s \left|\frac{ds}{d\Theta}\right|. \tag{69}$$

The absolute value takes into account the fact that as s increases, the particle is further from the target, so that the scattering angle Θ decreases. We now evaluate the three factors in (69). They are

i. $\frac{1}{\sin\Theta} = \frac{1}{2\sin\frac{\Theta}{2}\cos\frac{\Theta}{2}}$.

ii. $s = \dfrac{|k|}{2E\tan\frac{\Theta}{2}} = \dfrac{|k|\cos\frac{\Theta}{2}}{2E\sin\frac{\Theta}{2}}$.

iii. $\left|\dfrac{ds(\Theta)}{d\Theta}\right| = \dfrac{|k|}{4E\sin^2\frac{\Theta}{2}}$.

Multiplying these factors gives the answer (68). Thus to find the cross section, we need to check (i–iii).

The identity (i) is the double angle formula for $\sin 2\psi = 2\sin\psi\,\cos\psi$.

To check (ii), recall that $\cos\theta_{\max} = \frac{1}{\epsilon}$, so $\sin\theta_{\max} = \sqrt{1-\cos^2\theta_{\max}} = \sqrt{\frac{\epsilon^2-1}{\epsilon^2}}$. Hence

$$\tan\theta_{\max} = \frac{\sin\theta_{\max}}{\cos\theta_{\max}} = \frac{\frac{1}{\epsilon}\sqrt{\epsilon^2-1}}{\frac{1}{\epsilon}} = \sqrt{\epsilon^2-1} = \sqrt{\frac{2E}{\mu k^2}}\, L\ . \tag{70}$$

One can compute L for the incoming particle where asymptotically the energy is purely kinetic. Then $L = |\vec{r}\times\vec{p}| = sp_{\text{incoming}} = s\sqrt{2\mu E}$. Hence we can rewrite (70) as

$$\tan\theta_{\max} = s\sqrt{\frac{2E}{\mu k^2}}\sqrt{2\mu E} = s\frac{2E}{|k|}\ . \tag{71}$$

From the diagram we infer that $\Theta + 2\theta_{\max} = \pi$ or

$$\boxed{\theta_{\max} + \frac{\Theta}{2} = \frac{\pi}{2}}\ , \tag{72}$$

so $\sin\theta_{\max} = \cos\frac{\Theta}{2}$ and $\cos\theta_{\max} = \sin\frac{\Theta}{2}$. Therefore (71) can be written

$$\frac{1}{\tan\frac{\Theta}{2}} = \tan\theta_{\max} = s\frac{2E}{|k|}\ . \tag{73}$$

This is the identity (ii).

In order to derive identity (iii), note that

$$\frac{d}{du}\frac{1}{\tan u} = \frac{d}{du}\frac{\cos u}{\sin u} = -\frac{\sin u}{\sin u} - \frac{\cos^2 u}{\sin^2 u} = -\frac{\sin^2 u + \cos^2 u}{\sin^2 u} = -\frac{1}{\sin^2 u}\ . \tag{74}$$

Thus

$$\frac{ds(\Theta)}{d\Theta} = \frac{d}{d\Theta}\frac{|k|}{2E\tan\frac{\Theta}{2}} = -\frac{|k|}{4E\sin^2\frac{\Theta}{2}}\ .$$

Taking the absolute value gives the answer.

A1 Spherical Coordinates in 3 Dimensions

In the following appendices, we describe some standard mathematical notions used in these notes. We start with spherical polar coordinates in three dimensions. We need this picture to analyze the sphere around the target. The bottom line is that we want to write the volume element in three dimensions as $r^2\,dr\,d\Omega$, where r denotes a radial coordinate (radius of the sphere) and Ω denotes the solid angle–namely the fraction of the surface area of the sphere.

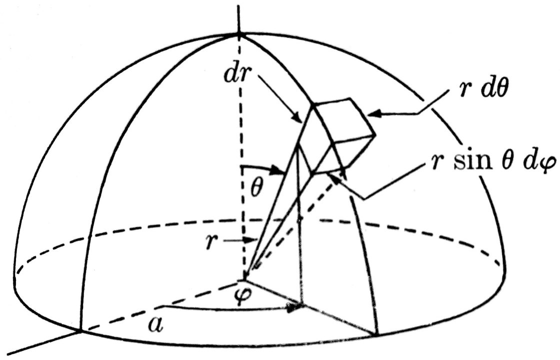

Figure 7: Solid angle Ω in spherical polar coordinates is parameterized by θ and φ.

The usual parameterization for spherical polar coordinates (r, θ, φ) in terms of Cartesian coordinates (x_1, x_2, x_3) is

$$x_1 = r \sin\theta \cos\varphi \;, \quad x_2 = r\sin\theta\sin\varphi \;, \quad x_3 = r\cos\theta \;. \tag{75}$$

Here θ is denotes the declination angle, and φ is the azimuthal angle. One can describe this in the following diagram. An increment of solid angle $d\Omega = \sin\theta d\theta d\varphi$ is illustrated in the diagram, as is the element of volume $dV = r^2 dr d\Omega$.

Alternatively one can find the volume element from the matrix of partial derivatives from one coordinate system as a function of the other. This matrix is central in geometry, and has been studied two hundred years; it generally called the "Jacobian" after the nineteenth century mathematician Carl Gustav Jacobi. In general, when one transforms from one coordinate system to another, the determinant $\det \widetilde{J}$ determines the transformation of the volume element. This generalizes the one dimensional formula for the transformation $x = x(\theta)$, and the relation $dx = \frac{dx}{d\theta} d\theta$ for the transformation of increments.

For the transformation from Cartesian to spherical coordinates, one denotes the Jacobian matrix as $\widetilde{J} = \frac{\partial(x_1,x_2,x_3)}{\partial(r,\theta,\varphi)}$, which just means that

$$\widetilde{J} = \frac{\partial(x_1,x_2,x_3)}{\partial(r,\theta,\varphi)} = \begin{pmatrix} \frac{\partial x_1}{\partial r} & \frac{\partial x_2}{\partial r} & \frac{\partial x_3}{\partial r} \\ \frac{\partial x_1}{\partial \theta} & \frac{\partial x_2}{\partial \theta} & \frac{\partial x_3}{\partial \theta} \\ \frac{\partial x_1}{\partial \varphi} & \frac{\partial x_2}{\partial \varphi} & \frac{\partial x_3}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\varphi & \sin\theta\sin\varphi & \cos\theta \\ r\cos\theta\cos\varphi & r\cos\theta\sin\varphi & -r\sin\theta \\ -r\sin\theta\sin\varphi & r\sin\theta\cos\varphi & 0 \end{pmatrix} \;. \tag{76}$$

The determinant is $\det\widetilde{J} = r^2\sin\theta$. The volume element $dV = dx_1\,dx_2\,dx_3$ in Cartesian coordinates, when expressed in terms of spherical polar coordinates becomes

$$dV = dx_1\,dx_2\,dx_3 = (\det\widetilde{J})\, dr d\theta d\varphi = r^2 \sin\theta dr d\theta d\varphi \;. \tag{77}$$

Likewise

$$d\Omega = (\det \widetilde{J})\, d\theta d\varphi \Big|_{r=1} = \sin\theta \, d\theta d\varphi \; . \tag{78}$$

A2 Three Useful Vector Identitites:

Before proceeding, we mention three identities for vectors in 3 dimensional space. We only use very special cases of these, but we state the general relations as they are often quite useful. I give conceptual proofs of these relations here, but one should check that if one takes $\vec{A}$ to have components A_j for $j = 1, 2, 3$, one can derive the same result by considering the components, as I also do for the first identity.

Identity I: For any three vectors $\vec{A}, \vec{B}, \vec{C}$ one has

$$\boxed{\left(\vec{A} \times \vec{B}\right) \cdot \vec{C} = \vec{A} \cdot \left(\vec{B} \times \vec{C}\right)} \; . \tag{79}$$

Both expressions equal the volume of the solid spanned by the three vectors. In components, they both equal $\sum_{i,j,k} \epsilon_{ijk} A_i B_j C_k$.

Identity II: For any three vectors $\vec{A}, \vec{B}, \vec{C}$ one has

$$\boxed{\vec{A} \times (\vec{B} \times \vec{C}) = (\vec{A} \cdot \vec{C})\vec{B} - (\vec{A} \cdot \vec{B})\vec{C}} \tag{80}$$

The triple product must lie in the plane spanned by $\vec{B}$ and $\vec{C}$. The coefficients must be chosen so the overall expression is tri-linear. Thus there are 3×3 matrices T, S such that

$$\vec{A} \times (\vec{B} \times \vec{C}) = (\vec{A} \cdot T\vec{C})\, \vec{B} - (\vec{A} \cdot S\vec{B})\, \vec{C} \; .$$

But antisymmetry under the interchange of $\vec{B}$ and $\vec{C}$ means that $S = T$. Furthermore the result must rotate like a vector, so $T = \lambda I$, where λ is a number. One can determine $\lambda = 1$ by evaluating the special case: $\vec{A} = \vec{B} = \vec{e}_1$, $\vec{C} = -\vec{e}_3$ for which $\vec{A} \times (\vec{B} \times \vec{C}) = \vec{e}_3$.

Identity III: Any two vectors $\vec{A}$ and $\vec{B}$ satisfy

$$\boxed{(\vec{A} \times \vec{B})^2 = \vec{A}^{\,2}\, \vec{B}^{\,2} - (\vec{A} \cdot \vec{B})^2} \; . \tag{81}$$

One way to see this is to use the Identities I and II in that order, namely

$$(\vec{A} \times \vec{B}) \cdot (\vec{A} \times \vec{B}) = \vec{A} \cdot (\vec{B} \times (\vec{A} \times \vec{B})) = \vec{A} \cdot (\vec{B}^2 \vec{A} - (\vec{B} \cdot \vec{A})\, \vec{B}) = \vec{A}^{\,2}\, \vec{B}^{\,2} - (\vec{A} \cdot \vec{B})^2 \; .$$

Alternatively, one can use the well-known formulas for the length of $\vec{A}\times\vec{B}$ and $\vec{A}\cdot\vec{B}$. The length of $\vec{A}\times\vec{B}$ is $\left|\vec{A}\right|\left|\vec{B}\right||\sin\theta|$, where θ is the angle between the vectors. Thus $(\vec{A}\times\vec{B})^2 = \vec{A}^2\vec{B}^2\sin^2\theta$. The identity then follows from $\vec{A}\cdot\vec{B} = \left|\vec{A}\right|\left|\vec{B}\right|\cos\theta$, along with $\sin^2\theta = 1-\cos^2\theta$.

A3 Conic Sections & Cartesian Coordinates

In this section we review some properties of conic sections when expressed in Cartesian coordinates. We relate this to the equation of the orbits found in polar coordinates.

A3.1 The Ellipse

To simplify things, let us orient our coordinates so that the longer axis is in the direction of the first axis. There are two standard ways to parameterize an ellipse.

Parameterization in Cartesian coordinates: An ellipse is the curve defined by the Cartesian coordinates (x_1, x_2) that satisfy

$$\left(\frac{x_1}{a}\right)^2 + \left(\frac{x_2}{b}\right)^2 = 1\ . \tag{82}$$

Here $a \geqslant b$ are the lengths of the semi-major and semi-minor axes respectively. In case $a = b$, then the ellipse is a circle of radius a.

Parameterization in terms of the foci: The ellipse in Figure 8 is the locus of points for which the sum of the distances from the two foci to a point on the ellipse equals the constant $2a$.

We claim that:

$$\boxed{r\,(1+\epsilon\cos\theta) = a\,(1-\epsilon^2)}\ . \tag{83}$$

Thus in the case of our orbit (36), one has $a\,(1-\epsilon^2) = \frac{L^2}{\mu k}$. Taking into account our expression for the eccentricity (33), this means that for any elliptical orbit,

$$\boxed{a = -\frac{k}{2E}}\ . \tag{84}$$

We derive (83) by considering two vectors as $\vec{r}$ with length r and the vector $\vec{r}'$ with length r', then $r + r' = 2a$. Since $\vec{r}' = \vec{r} + 2a\epsilon\,\widehat{e}_1$, one has

$$\begin{aligned} r'^{\,2} &= (\vec{r} + 2a\epsilon\,\widehat{e}_1)^2 = r^2 + 4a^2\epsilon^2 + 4a\epsilon r\cos\theta \\ &= (2a-r)^2 = r^2 + 4a^2 - 4ar\ . \end{aligned} \tag{85}$$

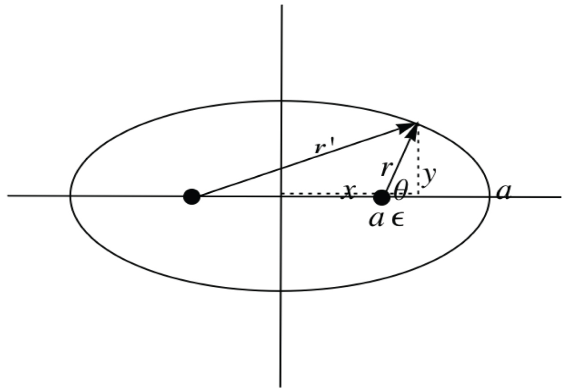

Figure 8: An ellipse with semi-major axis a and eccentricity ϵ.

Cancelling r^2 and dividing by $4a$ shows that (85) yields (83). Note that

$$r_{\text{min}}(1+\epsilon) = a(1-\epsilon^2) = a(1-\epsilon)(1+\epsilon) \; ,$$

so

$$r_{\text{min}} = a(1-\epsilon) \; . \tag{86}$$

Thus the focus on the right of Figure 8 is located at a distance $a - a(1-\epsilon) = a\epsilon$ from the center of the ellipse.

Relations between the two representations: Let us check that this parameterization yields exactly the same ellipse as the one in (83). We see from Figure 8 and the Pythagorean rule that

$$r = \sqrt{(x_1 - a\epsilon)^2 + x_2^2} \; , \quad \text{and} \quad r' = \sqrt{(x_1 + a\epsilon)^2 + x_2^2} \; . \tag{87}$$

Taking the special case $x_1 = 0$, $x_2 = b$ and using $r + r' = 2a$ shows that $2\sqrt{a^2\epsilon^2 + b^2} = 2a$. Hence

$$\boxed{b = a(1-\epsilon^2)^{1/2}} \; , \quad \text{or} \quad \boxed{\epsilon = \left(1 - \frac{b^2}{a^2}\right)^{1/2}} \; . \tag{88}$$

A3.2 The Parabola

In case that $E = 0$ or $\epsilon = 1$, the equation for the orbit is intermediate between an ellipse (for $\epsilon < 1$) and the unbound case for $\epsilon > 1$, which is a hyperbola, and which we analyze in separate notes. The intermediate orbit comes when $\epsilon = 1$, so the limit of elliptic orbits is

$r(1+\cos\theta) = \lim_{\epsilon\to 1} a(1-\epsilon^2)$. In this limit $E \to 0$, we know from (84) that the semi-major axis of the ellipse $a \to \infty$. But in the physics problem, the right side is just $\frac{L^2}{\mu k}$, which is finite. So in the limit $\epsilon = 1$, one must re-define the parameter on the right side of the equation. We choose $r_{\min} = a$. Then we take the equation of the limiting orbit to be

$$r(1+\cos\theta) = 2a \;, \tag{89}$$

and in the physics problem $2a = \frac{L^2}{\mu k}$. We claim that this is shown in Figure 9.

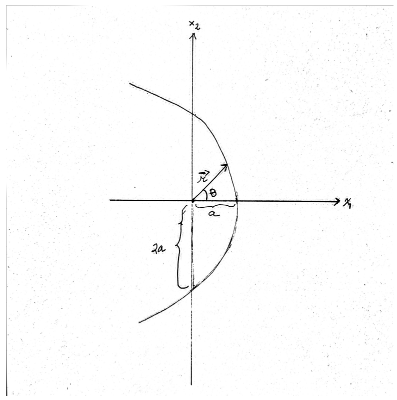

Figure 9: The parabola defined by Equation (89).

To check this, choose an origin for Cartesian coordinates at the focus, and let the angle θ be the angle between $\vec{r}$ and the x_1 axis, as in the figure. Then on the curve (89), $r = \sqrt{x_1^2 + x_2^2}$ and $x_1 = r\cos\theta$. So the equation (89) becomes $\sqrt{x_1^2+x_2^2} + x_1 = 2a$ or

$$\sqrt{x_1^2+x_2^2} = 2a - x_1 \;. \tag{90}$$

Squaring we see that

$$x_2^2 = 4a^2 - 4ax_1 \;, \tag{91}$$

or

$$\boxed{x_1 = a - \tfrac{1}{4a}\, x_2^2} \;. \tag{92}$$

This is the parabola (89). The curve crosses the x_1 axis (i.e. $x_2 = 0$) at the closest point to the focus, $r = x_1 = a$. This curve crosses the x_2 axis (i.e. $x_1 = 0$) when $r = 2a$ or $x_2 = \pm 2a$.

A3.3 The Hyperbola

The two branches of the hyperbola are shown in Figure 10, with the vector $\vec{r}$ pointing to the positive branch. In the physics problem, the positive branch corresponds to an attractive force, while the negative branch is used to describe a repulsive force.

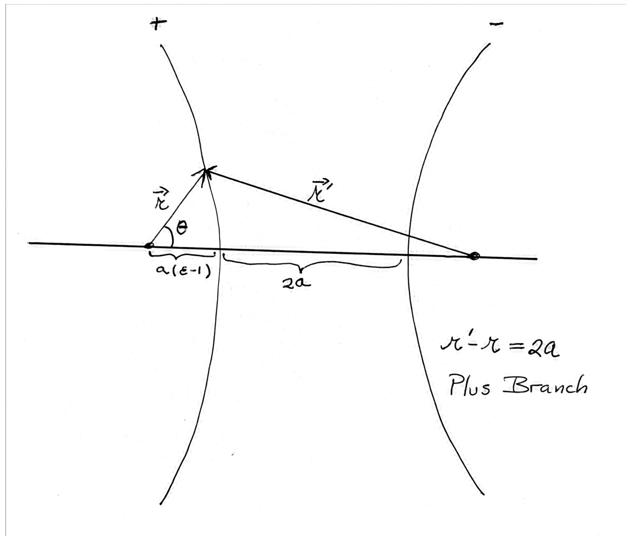

Figure 10: Positive Branch of a Hyperbola.

Positive Branch: Let's check that Figure 10 and the defining equation,

$$r' - r = 2a \ ,$$

corresponds to the equation of the orbit with eccentricity $1 < \epsilon$, namely

$$r(1 + \epsilon \cos\theta) = a(\epsilon^2 - 1) \ . \tag{93}$$

Let $\widehat{e}$ denote a unit vector oriented along the positive x_1 axis. Then

$$\vec{r} = 2a\epsilon\,\widehat{e} + \vec{r}\,' \ ,$$

so squaring $\vec{r} - 2a\epsilon\,\widehat{e}$ gives

$$r'^2 = r^2 + 4a^2\epsilon^2 - 4a\epsilon r\cos\theta = (r + 2a)^2 = r^2 + 4a^2 + 4a\epsilon r \ . \tag{94}$$

Simplifying gives

$$4a\epsilon r(1 + \epsilon\cos\theta) = 4a^2(\epsilon^2 - 1) \ , \tag{95}$$

which is (93). Furthermore, setting $\theta = 0$ shows that in the picture $r' - r$ is equal to

$$r' - r = (2a + a(\epsilon - 1)) - a(\epsilon - 1) = 2a \ , \tag{96}$$

as desired.

The maximum value of θ in (93) occurs for $1 + \epsilon \cos\theta_{\rm max} = 0$. As $\epsilon > 1$, this occurs for $\cos\theta_{\rm max} = -\frac{1}{\epsilon} < 0$, which means that $\theta_{\rm max}$ lies in the intervals where

$$\frac{\pi}{4} < \theta_{\rm max} < \frac{\pi}{2} \, . \tag{97}$$

Since $\cos -\theta = \cos\theta$, the hyperbola for $\theta < 0$ is the mirror through the $\theta = 0$ axis of the hyperbola for $\theta > 0$, the hyperbola goes off to infinity also for $\theta = -\theta_{\rm max}$.

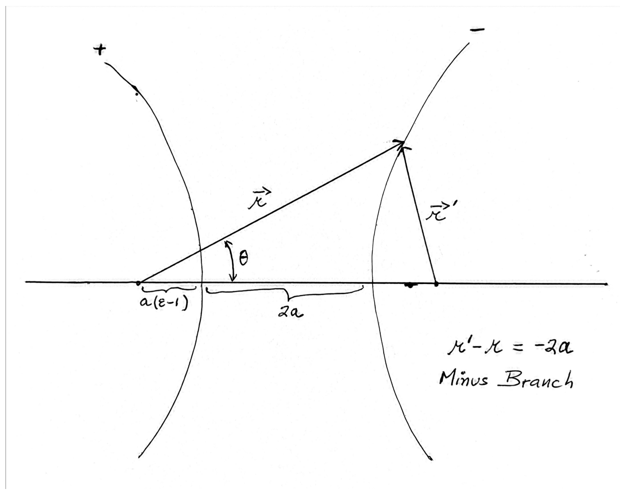

Figure 11: Negative Branch of a Hyperbola.

Negative Branch: The particle with position $\vec{r}$ on the negative branch of the hyperbola is shown in Figure 11, and this corresponds to a repulsive force problem. We check that Figure 11 and the defining equation $r' - r = -2a$ corresponds to the equation of the orbit,

$$r(-1 + \epsilon\cos\theta) = a(\epsilon^2 - 1) \, . \tag{98}$$

We use the relation $\vec{r} = 2a\epsilon\,\widehat{e} + \vec{r}\,'$, where $\widehat{e}$ again is a unit vector oriented along the x_1 axis. Calculating r'^2 one finds, similarly to (94), that

$$r'^2 = (\vec{r} - 2a\epsilon\widehat{e})^2 = r^2 + 4a^2\epsilon^2 - 4a\epsilon r\cos\theta \, . \tag{99}$$

However, in this case $r' - r = -2a$, so one has $r' = r - 2a$ and $r'^2 = r^2 + 4a^2 + 4ar$. Using (99),

$$r^2 + 4a^2\epsilon^2 - 4a\epsilon r\cos\theta = r^2 + 4a^2 - 4ar \, . \tag{100}$$

Now simplifying gives (98). As $r > r'$, we have on the negative branch that $r \to \infty$ for $\theta = \theta_{\mathrm{max}}$ or $\theta = -\theta_{\mathrm{max}}$, and

$$0 < \theta_{\mathrm{max}} < \frac{\pi}{4} \,. \tag{101}$$

Again $r \to \infty$, as θ in creases in magnitude from $\theta = 0$ to the limit $\theta \to \pm\theta_{\mathrm{max}}$.