

# Two- and Three-Dimensional Motion (Symon Chapter Three)

Physics A300\*

Spring 2006

## Contents

<b>1</b>	<b>Vector Analysis</b>	<b>3</b>
1.1	Definitions . . . . .	3
1.1.1	Notation . . . . .	3
1.2	Basic Vector Arithmetic . . . . .	3
1.2.1	Equality . . . . .	3
1.2.2	Magnitude . . . . .	4
1.2.3	Multiplication by a Scalar . . . . .	4
1.2.4	Addition of Vectors . . . . .	5
1.2.5	The Zero Vector . . . . .	6
1.3	Bases and Components . . . . .	6
1.4	The Scalar (Dot, Inner) Product . . . . .	6
1.5	The Vector (Cross) Product . . . . .	8
1.6	The Position Vector . . . . .	10
1.7	Differentiation and Integration . . . . .	12
1.7.1	Differentiation . . . . .	12
1.7.2	Integration . . . . .	12
<b>2</b>	<b>Vectors in Mechanics and Kinematics</b>	<b>13</b>
2.1	Force and Newton's Second Law . . . . .	13
2.2	Work and Energy . . . . .	14
<b>3</b>	<b>Non-Cartesian Coördinate Systems</b>	<b>15</b>
3.1	Plane Polar Coördinates . . . . .	15
3.1.1	Derivatives of Basis Vectors . . . . .	17
3.1.2	Consequences for Kinematics . . . . .	18
3.2	Cylindrical Coördinates . . . . .	19
3.3	Spherical Coördinates . . . . .	21

---

\*Copyright 2006, John T. Whelan, and all that

<b>4</b>	<b>Conservation Theorems in Two or More Dimensions</b>	<b>22</b>
4.1	Force and Momentum . . . . .	22
4.2	Work and Kinetic Energy . . . . .	23
4.2.1	Complications Arising from Potential Energy . . . . .	23
4.3	Conservative Forces . . . . .	24
<b>5</b>	<b>Fundamentals of Vector Calculus</b>	<b>27</b>
5.1	The Gradient . . . . .	27
5.2	Divergence, Curl, and All That . . . . .	28
5.3	Practical Calculation of Potential Energy . . . . .	29
5.4	Application of Potential Energy . . . . .	30
<b>6</b>	<b>Rotational Motion</b>	<b>32</b>
6.1	Torque . . . . .	32
6.2	Angular Momentum . . . . .	33
6.3	Relationship of Angular Momentum and Torque . . . . .	33
<b>7</b>	<b>Central Force Motion</b>	<b>34</b>
7.1	Conservation of Energy . . . . .	35
7.2	Conservation of Angular Momentum . . . . .	35
7.3	The Nature of the Differential Equations . . . . .	36
7.4	Effective Potential . . . . .	38
7.5	Finding the Trajectory . . . . .	39
7.6	Kepler's Second Law . . . . .	39
<b>8</b>	<b>Inverse-Square-Law Forces</b>	<b>40</b>
8.1	Finding the Turning Points . . . . .	40
8.2	Finding the Trajectory . . . . .	42
8.3	Orbits in an Inverse-Square Force Field . . . . .	43
8.4	Some notes for mechanics by Brans, March 24, 2006 . . . . .	44
8.4.1	Effective potential and "centrifugal force" . . . . .	44
8.4.2	Turning points . . . . .	45
8.4.3	Orbit geometry for Newtonian gravity. . . . .	45
8.4.4	Some elliptical notes on ellipses. . . . .	46
<b>A</b>	<b>Appendix: Correspondence to Class Lectures</b>	<b>47</b>

Tuesday, February 21, 2006

# 1 Vector Analysis

## 1.1 Definitions

There are numerous definitions of a vector, including

- An arrow in space
- Something defined by its magnitude and direction
- An ordered list of numbers
- A directional derivative

We'll start from the physical/geometrical definition in terms of magnitude and direction.

### 1.1.1 Notation

When talking about a vector, we will write it with an arrow, like so:  $\vec{A}$ . Symon uses boldface, like so:  $\mathbf{A}$ . To avoid confusion, it's very important never to accidentally drop the vector symbol from a vector, since the same letter without a vector symbol may mean something entirely different.

Note that vectors give us additional grammatical rules. Just as you can't add a length to a force, or compare a mass to a time, it's completely meaningless to try to say that a vector is equal to something that's not a vector. So it's okay to say  $\vec{A} = \vec{B}$  or  $a = b$ , but something like " $\vec{A} = a$ " just makes no sense.

## 1.2 Basic Vector Arithmetic

### 1.2.1 Equality

We think of a vector as being defined solely by its magnitude and direction, so if you've got two different vectors on a diagram in different places, you're free to "slide" them around without changing their directions to compare them. So in the diagram below  $\vec{A} = \vec{B}$ :



### 1.2.2 Magnitude

The magnitude or “length” of a vector is a positive number, and is *not* a vector. We write it with an absolute value sign, like so:  $|\vec{A}|$ . Note that the magnitude of a vector will not in general have units of length. The force  $\vec{F}$ , momentum  $\vec{p}$ , and acceleration  $\vec{a}$  are all vectors.

### 1.2.3 Multiplication by a Scalar

The simplest thing you can do to a vector is multiply it by a number, which is also called a scalar. (There are more precise definitions of a scalar, specifically that it should be something which doesn’t depend on the choice of coordinate system, but for the time being it’s fine to think of it as a number.)

It’s convenient to consider one special case first, namely multiplication by  $-1$ . In this case, we define the result to be a new vector with the same magnitude, pointing in the opposite direction:



Note that

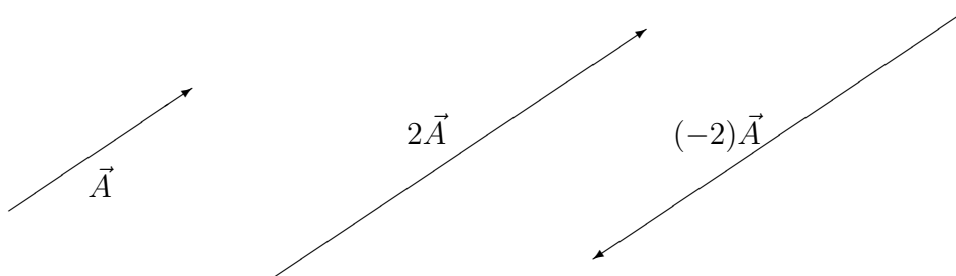
$$|-\vec{A}| = |\vec{A}| \quad (1.1)$$

Now, we can consider multiplication by a general scalar, which is defined as follows:  $c\vec{A}$  is a vector whose magnitude is

$$|c\vec{A}| = |c| |\vec{A}| \quad (1.2)$$

(note that we need the absolute value sign, to make sure the magnitude of  $c\vec{A}$  is not negative) and which points in the same direction as  $\vec{A}$  if  $c > 0$  and in the opposite direction if  $c < 0$ .

Here are some examples:

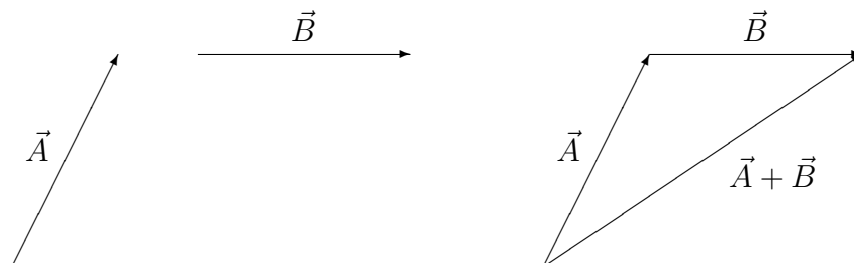


Multiplication of a scalar by a vector behaves in exactly the way you imagine it would; in particular it obeys the associative property:

$$b(c\vec{A}) = (bc)\vec{A} \quad (1.3)$$

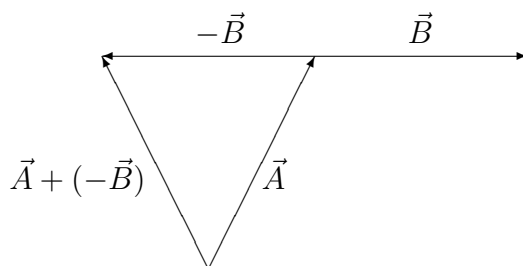
### 1.2.4 Addition of Vectors

The geometrical definition of the sum of two vectors follows from the idea of a vector as an arrow. To get the effect of  $\vec{A}$  plus  $\vec{B}$ , you just line the arrows up, putting the tail of  $\vec{B}$  at the head of  $\vec{A}$ , and their sum  $\vec{A} + \vec{B}$  goes from the tail of  $\vec{A}$  to the head of  $\vec{B}$ :



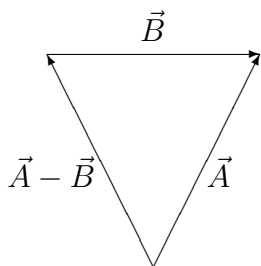
There are two reasonable ways to define subtraction of one vector from another. One is as multiplication by  $-1$  followed by addition:

$$\vec{A} - \vec{B} = \vec{A} + (-\vec{B}) \quad (1.4)$$



Alternatively, we can note that  $\vec{A} - \vec{B}$  is the vector to which you would add  $\vec{B}$  and be left with  $\vec{A}$ :

$$(\vec{A} - \vec{B}) + \vec{B} = \vec{A} \quad (1.5)$$



Both definitions are of course equivalent, and give the same result for  $\vec{A} - \vec{B}$ .

Note that with vector addition, as with ordinary addition, the quantities being added must have the same units. It makes no sense to write  $\vec{F} + \vec{p}$  if  $\vec{F}$  is a force and  $\vec{p}$  is a momentum.

Vector addition, together with multiplication by a scalar, obeys all of the properties you'd expect. These are collected in equations (3.5–3.8) of Symon.

Note that it makes no sense to try to add a vector to a scalar, just as it makes no sense to say that a vector equals a scalar. An expression like  $\vec{A} + a$  is meaningless.

### 1.2.5 The Zero Vector

Geometrically, if a vector has zero magnitude, its direction is undefined. (What direction does an arrow of zero length point in?) This means that there is only one vector with zero magnitude, which we call the zero vector  $\vec{0}$ . When a vector quantity vanishes, it is equal to this vector, e.g.,

$$\vec{A} - \vec{A} = \vec{0} \quad (1.6)$$

See also the footnote on Symon, page 79.

## 1.3 Bases and Components

These geometrical definitions are all well and good, but they can only go so far. Eventually we want to be able to do vector calculations algebraically, without having to draw a picture every time. To do this, we start with a Cartesian  $(x, y, z)$  coordinate system, and define unit vectors  $\hat{x}$ ,  $\hat{y}$  and  $\hat{z}$  parallel to the three coordinate axes. We write a hat over them, rather than an arrow, to emphasize that they're unit vectors, i.e., of magnitude one:

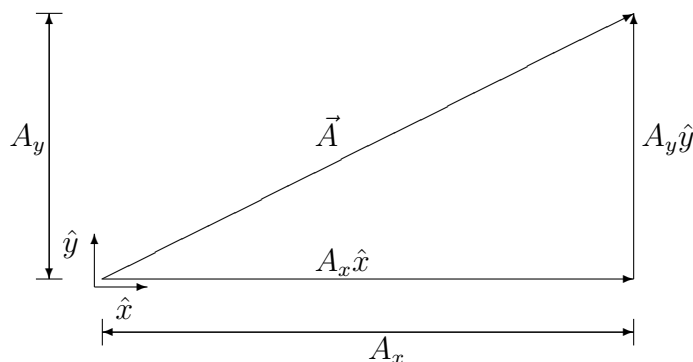
$$|\hat{x}| = |\hat{y}| = |\hat{z}| = 1 \quad (1.7)$$

Note that this means that the unit vectors  $\hat{x}$ ,  $\hat{y}$ , and  $\hat{z}$ , like all unit vectors, are dimensionless, and in particular that they do *not* have units of length like the coordinates  $x$ ,  $y$ , and  $z$ .

Given these unit vectors, one can project *any* vector  $\vec{A}$  onto the three coordinate axes, defining the projections as  $A_x$ ,  $A_y$ , and  $A_z$ . These are all numbers (*not* vectors) with units equal to those of the original vector  $\vec{A}$ . Geometrically, the law of vector addition makes it relatively easy to see that

$$\vec{A} = A_x \hat{x} + A_y \hat{y} + A_z \hat{z} \quad (1.8)$$

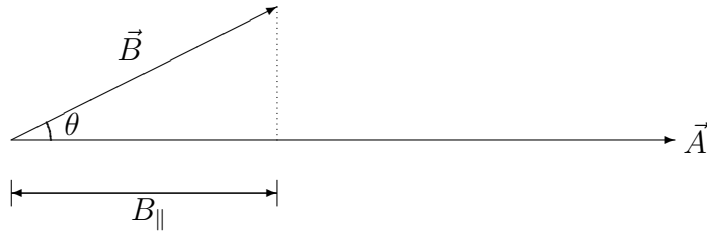
We illustrate this in two dimensions:



## 1.4 The Scalar (Dot, Inner) Product

We've defined how to add vectors and how to multiply a vector by a scalar. There turn out to be several different sensible ways to define the product of one vector with another, which have names corresponding to the object which results from the multiplication.

The simplest of these is the scalar product, also known as the dot product. Geometrically, it's the magnitude of one vector (e.g.,  $|\vec{A}|$ ) times the component of the second parallel to the first  $B_{\parallel}$ :



Using the right triangle whose hypotenuse is the length of  $\vec{B}$ , whose opening angle is  $\theta$ , and whose adjacent side has length  $B_{\parallel}$ , we see

$$\cos \theta = \frac{B_{\parallel}}{|\vec{B}|} \quad (1.9)$$

i.e.,

$$B_{\parallel} = |\vec{B}| \cos \theta \quad (1.10)$$

So the dot product is

$$\vec{A} \cdot \vec{B} := |\vec{A}| B_{\parallel} = |\vec{A}| |\vec{B}| \cos \theta \quad (1.11)$$

Note that

1. If the angle between the two vectors is more than  $\pi/2$  (i.e.,  $90^\circ$ ), the cosine, and thus the dot product, is negative.
2. The dot product is symmetric, i.e.,

$$\vec{A} \cdot \vec{B} = \vec{B} \cdot \vec{A} \quad (1.12)$$

You might worry that interchanging the role of  $\vec{A}$  and  $\vec{B}$  changes the sign of  $\theta$ , but cosine is an even function

$$\cos(-\theta) = \cos \theta \quad (1.13)$$

so the sign of the dot product doesn't change.

3. If  $\vec{A} \parallel \vec{B}$ , then  $\theta = 0$  and  $\vec{A} \cdot \vec{B} = |\vec{A}| |\vec{B}|$
4. If  $\vec{A} \perp \vec{B}$ , then  $\theta = \pi/2$  and  $\vec{A} \cdot \vec{B} = 0$

The dot product obeys some sensible identities such as commutation with scalar multiplication and distribution over vector addition, which appear as Symon's equations (3.16) and (3.17).

It's a simple matter of geometry (exercise!) to work out the dot products of the basis vectors:

$$\begin{aligned} \hat{x} \cdot \hat{x} &= 1 & \hat{x} \cdot \hat{y} &= 0 & \hat{x} \cdot \hat{z} &= 0 \\ \hat{y} \cdot \hat{x} &= 0 & \hat{y} \cdot \hat{y} &= 1 & \hat{y} \cdot \hat{z} &= 0 \\ \hat{z} \cdot \hat{x} &= 0 & \hat{z} \cdot \hat{y} &= 0 & \hat{z} \cdot \hat{z} &= 1 \end{aligned} \quad (1.14)$$

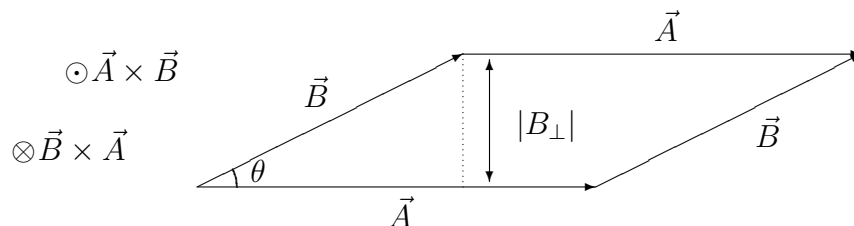
from which you can then show

$$\vec{A} \cdot \vec{B} = (A_x \hat{x} + A_y \hat{y} + A_z \hat{z}) \cdot (B_x \hat{x} + B_y \hat{y} + B_z \hat{z}) = A_x B_x + A_y B_y + A_z B_z \quad (1.15)$$

Which is sometimes treated as the algebraic definition of the dot product.

## 1.5 The Vector (Cross) Product

The scalar (dot) product is naturally defined in any number of dimensions. In three dimensions, it's also possible to define a vector product of two vectors, also known as the cross product. As a vector, it has a magnitude and a direction. Its magnitude is defined as the area of a parallelogram with the vectors  $\vec{A}$  and  $\vec{B}$  for sides:



Recalling that the area of a parallelogram is the length of the base (here  $|\vec{A}|$ ) times the height (here  $|B_{\perp}|$ , the component of  $\vec{B}$  perpendicular to  $\vec{A}$ ), and noting by trigonometry that

$$\sin \theta = \frac{|B_{\perp}|}{|\vec{B}|} \quad (1.16)$$

we get the magnitude

$$|\vec{A} \times \vec{B}| := |\vec{A}| |B_{\perp}| = |\vec{A}| |\vec{B}| \sin \theta \quad (1.17)$$

The direction of this vector is perpendicular to the plane defined by the two vectors, which is why the definition only works in three dimensions. We still have to specify whether it's “up” or “down” from that plane, i.e., “out of” or “into” the page in the diagram above. This is done with the right-hand rule, which says that if you point your right hand parallel to the first vector and then curl your fingers in the direction of the second vector, your thumb will point in the direction of the third, which in the diagram above is out of the page. Note that this means that if we switch the order of the vectors in a cross product, we reverse the direction of the result:

$$\vec{B} \times \vec{A} = -\vec{A} \times \vec{B} \quad (1.18)$$

We therefore say that the cross product is *antisymmetric*. Note also that

1. If  $\vec{A} \parallel \vec{B}$ , then  $\theta = 0$  and  $|\vec{A} \times \vec{B}| = 0$ , which means  $\vec{A} \times \vec{B} = \vec{0}$ .

2. If  $\vec{A} \perp \vec{B}$ , then  $\theta = \pi/2$  and  $|\vec{A} \times \vec{B}| = |\vec{A}| |\vec{B}|$

To get the cross product in component form, we first work out the cross products of the basis vectors (exercise!):

$$\begin{aligned} \hat{x} \times \hat{x} &= \vec{0} & \hat{x} \times \hat{y} &= \hat{z} & \hat{x} \times \hat{z} &= -\hat{y} \\ \hat{y} \times \hat{x} &= -\hat{z} & \hat{y} \times \hat{y} &= \vec{0} & \hat{y} \times \hat{z} &= \hat{x} \\ \hat{z} \times \hat{x} &= \hat{y} & \hat{z} \times \hat{y} &= -\hat{x} & \hat{z} \times \hat{z} &= \vec{0} \end{aligned} \quad (1.19)$$



and then we can find (exercise!)

$$\begin{aligned}\vec{A} \times \vec{B} &= (A_x \hat{x} + A_y \hat{y} + A_z \hat{z}) \times (B_x \hat{x} + B_y \hat{y} + B_z \hat{z}) \\ &= (A_y B_z - A_z B_y) \hat{x} + (A_z B_x - A_x B_z) \hat{y} + (A_x B_y - A_y B_x) \hat{z} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} \end{aligned} \quad (1.20)$$

where in the last step we have used the notation for the determinant of a matrix to write the answer compactly.

Two other useful identities (which we'll state here, but which can be proven) concern “triple” products:

$$\vec{A} \cdot (\vec{B} \times \vec{C}) = (\vec{A} \times \vec{B}) \cdot \vec{C} = \begin{vmatrix} A_x & A_y & A_z \\ B_x & B_y & B_z \\ C_x & C_y & C_z \end{vmatrix} \quad (1.21)$$

$$\vec{A} \times (\vec{B} \times \vec{C}) = (\vec{A} \cdot \vec{C})\vec{B} - (\vec{A} \cdot \vec{B})\vec{C} \neq (\vec{A} \times \vec{B}) \times \vec{C} \quad (1.22)$$

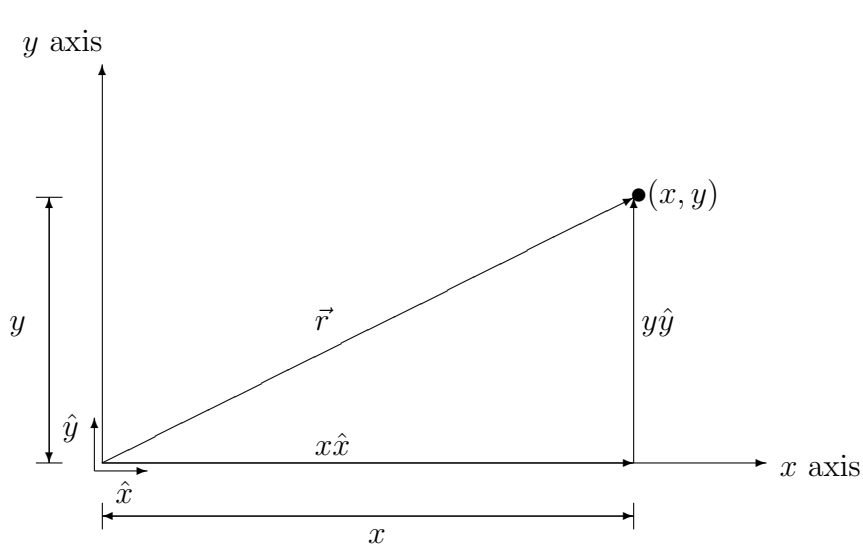
Friday, February 24, 2006

## 1.6 The Position Vector

If we have a Cartesian  $(x, y, z)$  coordinate system, we can define a vector of particular interest. For any point in space we can draw a vector from the origin of coordinates to that point; it's not too hard to see that the  $x$ -component of the vector will be the  $x$ -coordinate, the  $y$ -component of the vector will be the  $y$ -coordinate etc. We call this vector the *position vector* for that point:

$$\vec{r} = x\hat{x} + y\hat{y} + z\hat{z} \quad (1.23)$$

We can illustrate this in two dimensions:

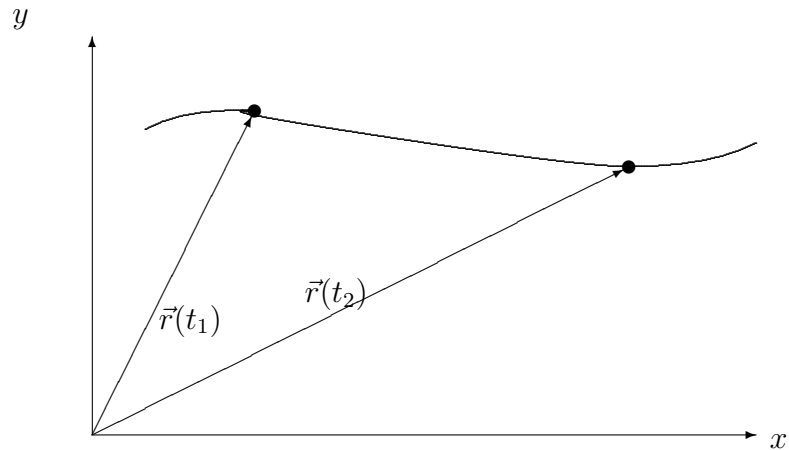


Each point has its own position vector. Note also that the position vector depends on the location of the origin of our coordinate system.

We can describe the trajectory of a particle by specifying its location (i.e., its Cartesian coordinates or equivalently its position vector) at each moment in time:

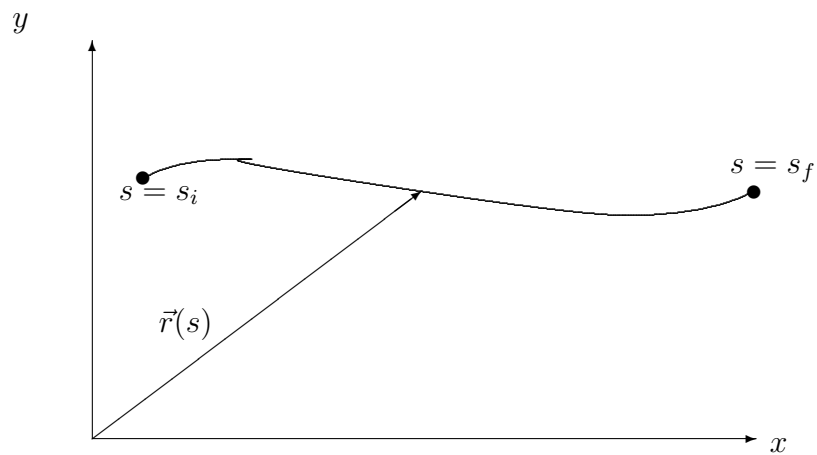
$$\vec{r}(t) = x(t)\hat{x} + y(t)\hat{y} + z(t)\hat{z} \quad (1.24)$$

We can illustrate this in two dimensions:



In fact, you can describe any curve through space as a position vector which is a function of some parameter  $s$ :

$$\vec{r}(s) = x(s)\hat{x} + y(s)\hat{y} + z(s)\hat{z} \quad (1.25)$$



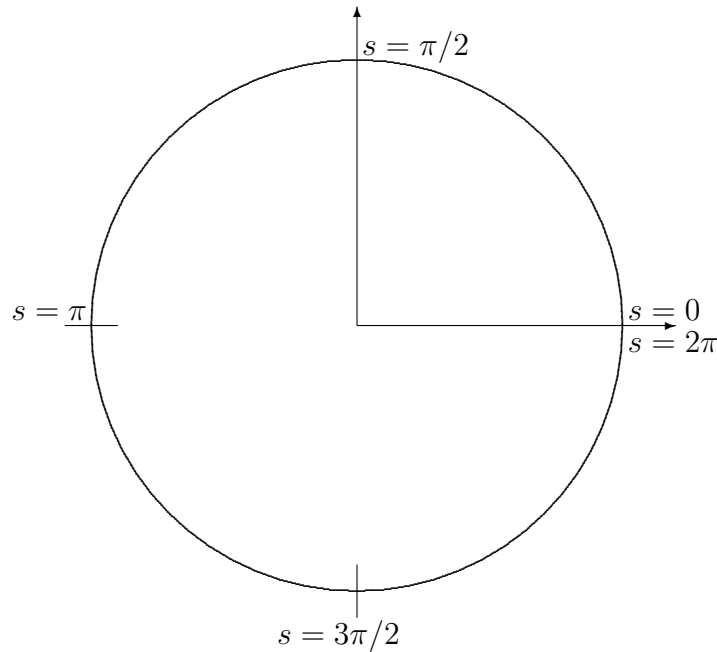
For example, a circle of radius  $R$  in the  $xy$ -plane centered at the origin is defined by

$$x^2 + y^2 = R^2 \quad (1.26)$$

A path which goes counter-clockwise around that circle starting at  $(x, y) = (R, 0)$  can be parametrized as

$$x = R \cos s \quad (1.27a)$$

$$y = R \sin s \quad (1.27b)$$



## 1.7 Differentiation and Integration

### 1.7.1 Differentiation

If  $\vec{A}(s)$  is a function of some parameter  $s$ , we can define a derivative in the usual way, as the limit of the ratio of infinitesimal changes:

$$\frac{d\vec{A}}{ds} := \lim_{\Delta s \rightarrow 0} \frac{\vec{A}(s + \Delta s) - \vec{A}(s)}{\Delta s} = \frac{dA_x}{ds} \hat{x} + \frac{dA_y}{ds} \hat{y} + \frac{dA_z}{ds} \hat{z} \quad (1.28)$$

This is because subtracting one vector from another is well defined, as is dividing a vector by a scalar.

As an example, consider the derivative of the position vector  $\vec{r}(t)$  with respect to time:

$$\frac{d\vec{r}}{dt} = \frac{dx}{dt} \hat{x} + \frac{dy}{dt} \hat{y} + \frac{dz}{dt} \hat{z} = \vec{v} \quad (1.29)$$

This is just the velocity vector.

This operation of differentiation obeys the usual sum, product and chain rules (see Symon's (3.54–3.57)), for example

$$\frac{d}{ds}(\vec{A} \times \vec{B}) = \frac{d\vec{A}}{ds} \times \vec{B} + \vec{A} \times \frac{d\vec{B}}{ds} \quad (1.30)$$

Note that because the “product” in this product rule is a cross product, we have to be careful to preserve the order of the vectors  $\vec{A}$  and  $\vec{B}$  in the last term of the answer!

### 1.7.2 Integration

Now things get a little more interesting. One quantity we can define is the integral along some path in space of a *vector field*.

A vector field (also known as a vector point function) is a function which defines a (possibly different) vector  $\vec{A}(\vec{r})$  for every point in space (represented by its position vector  $\vec{r}$ ).

We can consider a curve  $\mathcal{C}$  defined by  $\vec{r}(s)$  as  $s$  goes from  $s_i$  to  $s_f$ , and define the line integral

$$\oint_{\mathcal{C}} \vec{A} \cdot d\vec{r} \quad (1.31)$$

What does this mean? First we have to ask what is the differential  $d\vec{r}$ . Well, given the function  $\vec{r}(s)$  we can use the chain rule to write

$$d\vec{r} = \frac{d\vec{r}}{ds} ds \quad (1.32)$$

and then write

$$\oint_{\mathcal{C}} \vec{A} \cdot d\vec{r} = \int_{s_i}^{s_f} \left( \vec{A} \cdot \frac{d\vec{r}}{ds} \right) ds \quad (1.33)$$

Note that this depends in general on the curve  $\mathcal{C}$  and not just on the endpoints  $\vec{r}(s_i)$  and  $\vec{r}(s_f)$ , which is why we don't use some notation like  $\int_{\vec{r}_i}^{\vec{r}_f} (\dots) \cdot d\vec{r}$ . However, it turns out not to depend on the parametrization of the path, which can be shown by changing variables in the integral and using the chain rule.

## 2 Vectors in Mechanics and Kinematics

### 2.1 Force and Newton's Second Law

We've now defined enough vector quantities to start applying them to physics. For example, Newton's second law reads in one dimension

$$F_{\text{net}} = ma = m \frac{d^2x}{dt^2} = \frac{dp}{dt} \quad (2.1)$$

The three-dimensional version is

$$\vec{F}_{\text{net}} = m\vec{a} = m \frac{d^2\vec{r}}{dt^2} = \frac{d\vec{p}}{dt} \quad (2.2)$$

Where now the force, acceleration and momentum are all vectors. To convert the equation, we've just put vector signs on everything, except that the trajectory is now  $\vec{r}(t)$  rather than  $x(t)$ . The velocity is

$$\vec{v} = \frac{d\vec{r}}{dt} \quad (2.3)$$

the derivative of a vector, and the other definitions of kinematics carry over as  $\vec{a} = \frac{d\vec{v}}{dt}$  and  $\vec{p} = m\vec{v}$ .

Can it really be as easy as all that, though? More or less, yes. Recall that in chapter one we had a version of Newton's 2nd law written out component by component:

$$F_x = m\ddot{x} \quad (2.4a)$$

$$F_y = m\ddot{y} \quad (2.4b)$$

$$F_z = m\ddot{z} \quad (2.4c)$$

This is equivalent to the single vector equation (2.2), written as

$$F_x \hat{x} + F_y \hat{y} + F_z \hat{z} = \vec{F} = m\ddot{\vec{r}} = m(\ddot{x} \hat{x} + \ddot{y} \hat{y} + \ddot{z} \hat{z}) \quad (2.5)$$

because two vectors are equal if and only if all their components are equal.

## 2.2 Work and Energy

Note that not every physical quantity becomes a vector when we go to three dimensions. For example, the one-dimensional kinetic energy

$$T = \frac{1}{2}mv^2 \quad (2.6)$$

becomes

$$T = \frac{1}{2}m\vec{v} \cdot \vec{v} \quad (2.7)$$

which is a scalar.

Similarly, our definition of the line integral of a vector allows us to convert the one-dimensional definition of work

$$W_{1 \rightarrow 2} = \int_{x_1}^{x_2} F dx \quad (2.8)$$

into a line integral

$$W_{1 \rightarrow 2} = \int_C \vec{F} \cdot d\vec{r} \quad (2.9)$$

which is also a scalar.

We'll return to the work-energy theorem, but first we want to consider another complication, the use of non-Cartesian coordinate systems and basis vectors.

Friday, March 3, 2006

### 3 Non-Cartesian Coördinate Systems

Sometimes it's useful to locate points in a plane by coördinates other than  $\{x, y, z\}$  and also useful to resolve vectors along a basis other than  $\{\hat{x}, \hat{y}, \hat{z}\}$ .

#### 3.1 Plane Polar Coördinates

To go into detail for the simplest example, consider the case of plane polar coördinates  $\{r, \phi\}$  in two dimensions, which are related to the standard Cartesian coördinates by

$$x = r \cos \phi \tag{3.1a}$$

$$y = r \sin \phi \tag{3.1b}$$

which can be inverted to give<sup>1</sup>

$$r = \sqrt{x^2 + y^2} \tag{3.3a}$$

$$\tan \phi = \frac{y}{x} \tag{3.3b}$$

In Cartesian coördinates, we define the basis vector  $\hat{x}$  to point parallel to the  $x$  axis, i.e., in the direction of increasing  $x$  at constant  $y$  (and  $z$ , in three dimensions). The basis vector  $\hat{y}$  points in the direction of increasing  $y$  (and  $z$ , in three dimensions). (See figure 3.1) If we use the “direction of increasing coördinate” definition rather than the “parallel to the axis” definition, we can actually apply it everywhere and still get the same  $\hat{x}$  and  $\hat{y}$ . In polar coördinates, there's no “ $r$  axis” or “ $\phi$  axis”, but we can, at any point, define a unit vector  $\hat{r}$  which points in the direction of increasing  $r$  at constant  $\phi$ , and  $\hat{\phi}$  which points in the direction of increasing  $\phi$  at constant  $r$ . But these vectors will point in different directions in different places, so each of them is actually a *vector field*.

We can relate the polar coördinate basis vectors associated with a point to the Cartesian basis vectors by considering the position vector  $\vec{r}$  for that point. This is a vector which points from the origin to the point in question, so by definition it points along a line of constant  $\phi$ , in the radial direction (increasing  $r$ ). Its length is the distance to the origin, which is just the polar coördinate  $r$ . So

$$\vec{r} = r \hat{r} \tag{3.4}$$

---

<sup>1</sup>We can't quite define  $\phi$  as  $\tan^{-1}(y/x)$  because the arctangent is defined to lie between  $-\pi/2$  and  $\pi/2$ , which would only represent the first and fourth quadrants of the  $xy$ -plane, i.e.,  $x \geq 0$ . The most concise way to define  $\phi$  is to use the “`atan2`” function from various computer languages, which can be defined (somewhat pedantically) by

$$\text{atan2}(y, x) = \begin{cases} \tan^{-1}(y/x) - \pi & x < 0 \text{ and } y < 0 \\ -\pi/2 & x = 0 \text{ and } y < 0 \\ \tan^{-1}(y/x) & x > 0 \\ \pi/2 & x = 0 \text{ and } y > 0 \\ \tan^{-1}(y/x) + \pi & x < 0 \text{ and } y \geq 0 \end{cases} \tag{3.2}$$

in which case  $\phi = \text{atan2}(y, x) \bmod 2\pi$ .

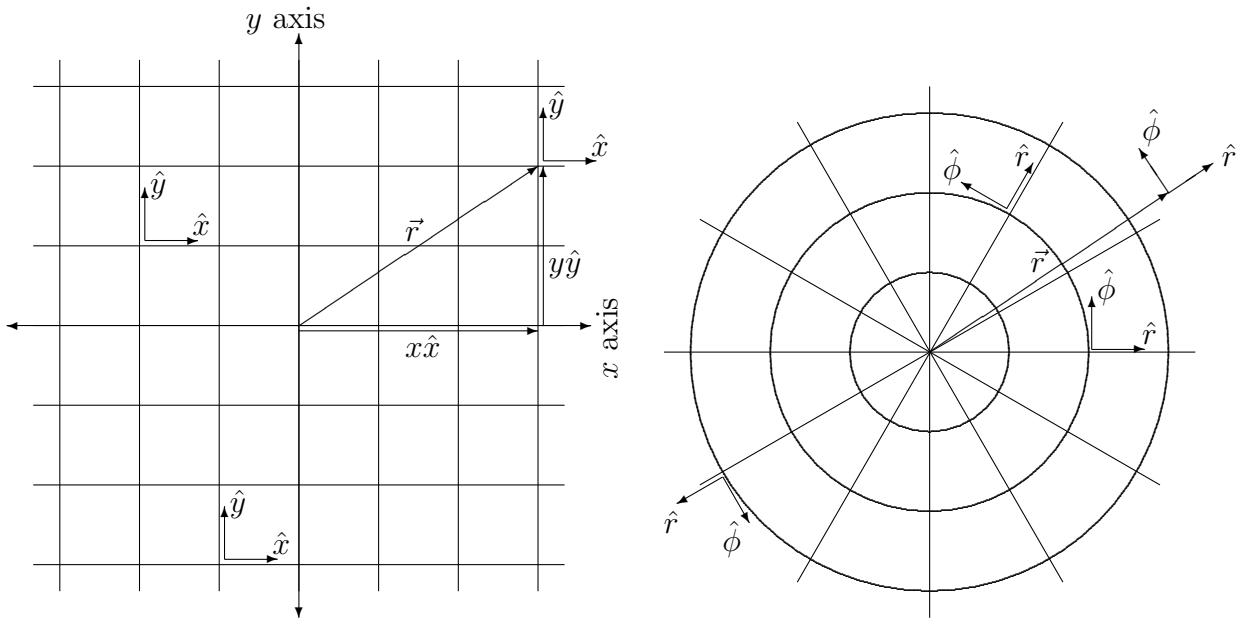


Figure 1: Constant-coördinate lines and basis vectors in Cartesian and polar coördinates.

(Note that there is *no* component in the  $\phi$  direction.) But we also know that

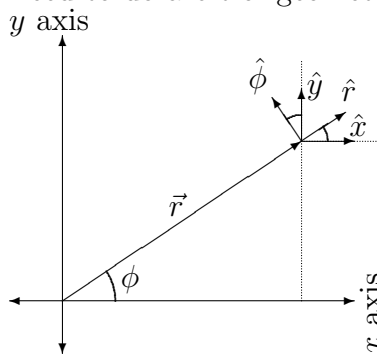
$$\vec{r} = x \hat{x} + y \hat{y} \tag{3.5}$$

which means we can solve for

$$\hat{r} = \frac{x}{r} \hat{x} + \frac{y}{r} \hat{y} = \hat{x} \cos \phi + \hat{y} \sin \phi \tag{3.6}$$

where we have substituted the coördinate transformations (3.1).

To find the basis vector  $\hat{\phi}$  we need to do a bit of geometry:



The marked angles are all  $\phi$ , so  $\hat{\phi}$  has an  $x$  component of  $-\sin \phi$  and a  $y$  component of  $\cos \phi$ . The components are thus

$$\hat{r} = \hat{x} \cos \phi + \hat{y} \sin \phi \tag{3.7a}$$

$$\hat{\phi} = -\hat{x} \sin \phi + \hat{y} \cos \phi \tag{3.7b}$$



Using the forms of the basis vectors in (3.7) one can check (exercise!) that the dot products of the basis vectors are

$$\begin{aligned} \hat{r} \cdot \hat{r} &= 1 & \hat{r} \cdot \hat{\phi} &= 0 \\ \hat{\phi} \cdot \hat{r} &= 0 & \hat{\phi} \cdot \hat{\phi} &= 1 \end{aligned} \quad (3.8)$$

The vectors  $\hat{r}$  and  $\hat{\phi}$  this form an *orthonormal basis*. “Ortho-” means the vectors are orthogonal or perpendicular and “normal” means they’re normalized, i.e., unit vectors.

### 3.1.1 Derivatives of Basis Vectors

Since the basis vectors  $\hat{r}$  and  $\hat{\phi}$  vary from point to point, it’s useful to calculate their partial derivatives with respect to the corresponding coördinates. We can do this by differentiating the expressions (3.7) and noting that  $\hat{x}$  and  $\hat{y}$  are constant to get

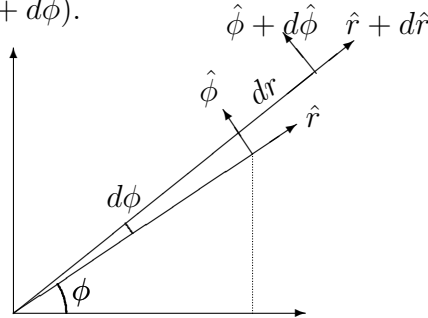
$$\left( \frac{\partial \hat{r}}{\partial r} \right)_{\phi} = \vec{0} \quad (3.9a)$$

$$\left( \frac{\partial \hat{\phi}}{\partial r} \right)_{\phi} = \vec{0} \quad (3.9b)$$

$$\left( \frac{\partial \hat{r}}{\partial \phi} \right)_r = -\hat{x} \sin \phi + \hat{y} \cos \phi = \hat{\phi} \quad (3.9c)$$

$$\left( \frac{\partial \hat{\phi}}{\partial \phi} \right)_r = -\hat{x} \cos \phi - \hat{y} \sin \phi = -\hat{r} \quad (3.9d)$$

We can also see the same thing geometrically if we consider the unit vectors at some point  $(r, \phi)$  and a nearby point  $(r + dr, \phi + d\phi)$ .



Resolving the basis vectors that are rotated by the infinitesimal angle  $d\phi$  into components along the unrotated basis vectors gives us

$$\hat{r} + d\hat{r} = \hat{r} \cos(d\phi) + \hat{\phi} \sin(d\phi) = \hat{r} + \hat{\phi} d\phi \quad (3.10a)$$

$$\hat{\phi} + d\hat{\phi} = \hat{\phi} \cos(d\phi) - \hat{r} \sin(d\phi) = \hat{\phi} - \hat{r} d\phi \quad (3.10b)$$

where we have used the small angle approximation to simplify the trig functions of the infinitesimal angle  $d\phi$ . From this we can read off

$$d\hat{r} = \hat{\phi} d\phi \quad (3.11a)$$

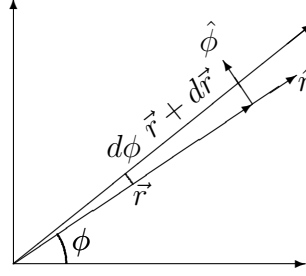
$$d\hat{\phi} = -\hat{r} d\phi \quad (3.11b)$$

which is another way of summarizing the partial derivatives (3.9). (See also Symon's Fig. 3.21.)

An important consequence of these derivatives is the differential of the position vector itself:

$$d\vec{r} = d(r \hat{r}) = dr \hat{r} + r d\hat{r} = dr \hat{r} + r(\dot{\phi} d\phi) = dr \hat{r} + (r d\phi) \hat{\phi} \quad (3.12)$$

We can also see this geometrically by drawing the vectors  $\vec{r}$  and  $\vec{r} + d\vec{r}$  and noting the components of  $d\vec{r}$ .



### 3.1.2 Consequences for Kinematics

The derivatives of the polar coordinate basis vectors allow us to write the velocity and acceleration in both Cartesian and polar coordinates:

$$\begin{aligned} \vec{v} &= \frac{d\vec{r}}{dt} = \frac{d}{dt}(x \hat{x} + y \hat{y}) = \underbrace{\dot{x}}_{v_x} \hat{x} + \underbrace{\dot{y}}_{v_y} \hat{y} \\ &= \frac{d}{dt}(r \hat{r}) = \dot{r} \hat{r} + r \frac{d\hat{r}}{dt} = \underbrace{\dot{r}}_{v_r} \hat{r} + \underbrace{r \dot{\phi}}_{v_\phi} \hat{\phi} \end{aligned} \quad (3.13)$$

where we have used the chain rule to describe the change in the unit vector  $\hat{r}$  as the particle moves from its position at time  $t$  to its position at time  $t + dt$ :

$$\frac{d\hat{r}}{dt} = \frac{\partial \hat{r}}{\partial r} \frac{dr}{dt} + \frac{\partial \hat{r}}{\partial \phi} \frac{d\phi}{dt} = \hat{\phi} \dot{\phi} \quad (3.14)$$

We could also obtain the same expression by dividing (3.13) by the infinitesimal time  $dt$  it takes to move from  $\vec{r}$  to  $\vec{r} + d\vec{r}$ .

Similarly, the acceleration can be written

$$\begin{aligned} \vec{a} &= \frac{d\vec{v}}{dt} = \frac{d}{dt}(\dot{r} \hat{r} + r \dot{\phi} \hat{\phi}) = \ddot{r} \hat{r} + \dot{r} \left[ \frac{\partial \hat{r}}{\partial r} \dot{r} + \frac{\partial \hat{r}}{\partial \phi} \dot{\phi} \right] + (\dot{r} \dot{\phi} + r \ddot{\phi}) \hat{\phi} + r \dot{\phi} \left[ \frac{\partial \hat{\phi}}{\partial r} \dot{r} + \frac{\partial \hat{\phi}}{\partial \phi} \dot{\phi} \right] \\ &= (\ddot{r} - r \dot{\phi}^2) \hat{r} + (r \ddot{\phi} + 2\dot{r} \dot{\phi}) \hat{\phi} \end{aligned} \quad (3.15)$$

As an example, consider the case of uniform circular motion at radius  $R$  and angular frequency  $\omega$ . In that case

$$r \equiv R \tag{3.16a}$$

$$\phi = \phi_0 + \omega t \tag{3.16b}$$

$$\dot{r} = 0 \tag{3.16c}$$

$$\dot{\phi} = \omega = \text{constant} \tag{3.16d}$$

and thus

$$\vec{v} = R\omega \hat{\phi} \tag{3.17a}$$

$$\vec{a} = -R\omega^2 \hat{r} \tag{3.17b}$$

## 3.2 Cylindrical Coördinates

We can take all our results for plane polar coördinates and use them as the basis for cylindrical coördinates in three dimensions. The coördinates are:

- $q$ , the distance of the point from the  $z$  axis;<sup>2</sup> note that like  $r$  in two or three dimensions,  $q$  is defined to be non-negative,  $0 \leq q < \infty$ .
- $\phi$ , an angle measured around the  $z$  axis; to cover the whole space, we have to let  $\phi$  take on any value in a  $2\pi$  range, e.g.,  $0 \leq \phi < 2\pi$ .
- $z$ , the height above the  $xy$ -plane, which can take any value  $-\infty < z < \infty$ .

They are related to three-dimensional Cartesian coördinates by

$$x = q \cos \phi \tag{3.18a}$$

$$y = q \sin \phi \tag{3.18b}$$

$$z = z \tag{3.18c}$$

and

$$q = \sqrt{x^2 + y^2} \tag{3.19a}$$

$$\phi = \text{atan2}(y, x) \pmod{2\pi} \tag{3.19b}$$

$$z = z \tag{3.19c}$$

We already know from our previous calculations (now applied in a plane of constant  $z$ ) that

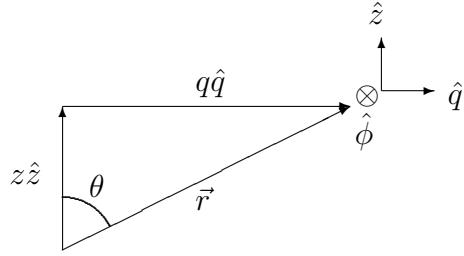
$$\hat{q} = \hat{x} \cos \phi + \hat{y} \sin \phi \tag{3.20a}$$

$$\hat{\phi} = -\hat{x} \sin \phi + \hat{y} \cos \phi \tag{3.20b}$$

$$\hat{z} = \hat{z} \tag{3.20c}$$

---

<sup>2</sup>We can't call this  $r$  because that is used in spherical coördinates for the distance from the origin  $r = |\vec{r}| = \sqrt{x^2 + y^2 + z^2}$ . Symon calls the cylindrical coördinate  $\rho$ , but we will need that letter later on for the density. Note that Griffiths, in *Introduction to Electrodynamics*, uses  $s$  for this coördinate.



By looking at the plane containing the position vector  $\vec{r}$  and the  $z$  axis we can see that

$$\vec{r} = q\hat{q} + z\hat{z} \quad (3.21)$$

and things pretty much proceed as before. In particular,

$$d\vec{r} = dq\hat{q} + (q\,d\phi)\hat{\phi} + z\hat{z} \quad (3.22)$$

Monday, March 6, 2006

### 3.3 Spherical Coördinates

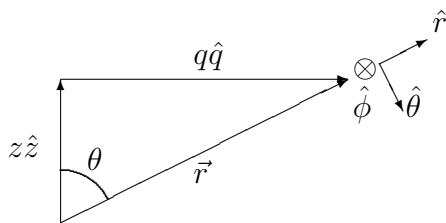
Spherical coördinates in three dimensions, like polar coördinates in two dimensions, are based around the distance from the origin

$$r = |\vec{r}| = \sqrt{x^2 + y^2 + z^2} \quad (3.23)$$

and the radial unit vector

$$\hat{r} = \frac{\vec{r}}{r} = \frac{x}{r} \hat{x} + \frac{y}{r} \hat{y} + \frac{z}{r} \hat{z} \quad (3.24)$$

The other two coördinates are  $\theta$  (the angle between the position vector  $\vec{r}$  and the  $z$  axis) and  $\phi$  (the same azimuthal angle used in cylindrical coördinates). The mathematical definitions of these angles and the corresponding basis vectors are most easily obtained by comparing spherical and cylindrical coördinates, looking at one plane at a time.



First, looking at the plane containing  $\vec{r}$  and  $\hat{z}$  we can work out

$$q = r \sin \theta \quad (3.25a)$$

$$z = r \cos \theta \quad (3.25b)$$

and

$$\hat{r} = \hat{q} \sin \theta + \hat{z} \cos \theta \quad (3.26a)$$

$$\hat{\theta} = \hat{q} \cos \theta - \hat{z} \sin \theta \quad (3.26b)$$

Looking in the  $xy$ -plane gives us the same relations (3.20a–3.20b) and (3.18a–3.18b); when we combine them with (3.25) and (3.26), we can relate the spherical and Cartesian systems as follows:

$$x = r \sin \theta \cos \phi \quad (3.27a)$$

$$y = r \sin \theta \sin \phi \quad (3.27b)$$

$$z = r \cos \theta \quad (3.27c)$$

and

$$\hat{r} = \hat{x} \sin \theta \cos \phi + \hat{y} \sin \theta \sin \phi + \hat{z} \cos \theta \quad (3.28a)$$

$$\hat{\theta} = \hat{x} \cos \theta \cos \phi + \hat{y} \cos \theta \sin \phi - \hat{z} \sin \theta \quad (3.28b)$$

$$\hat{\phi} = -\hat{x} \sin \phi + \hat{y} \cos \phi \quad (3.28c)$$

You can (and will on the homework) use these forms to verify that  $\{\hat{r}, \hat{\theta}, \hat{\phi}\}$  form an orthonormal basis ( $\hat{r} \cdot \hat{r} = 1$ ,  $\hat{r} \cdot \hat{\theta} = 0$ , etc) and to work out the derivatives

$$\begin{aligned} \frac{\partial \hat{r}}{\partial r} &= \vec{0} & \frac{\partial \hat{r}}{\partial \theta} &= \hat{\theta} & \frac{\partial \hat{r}}{\partial \phi} &= \hat{\phi} \sin \theta \\ \frac{\partial \hat{\theta}}{\partial r} &= \vec{0} & \frac{\partial \hat{\theta}}{\partial \theta} &= -\hat{r} & \frac{\partial \hat{\theta}}{\partial \phi} &= \hat{\phi} \cos \theta \\ \frac{\partial \hat{\phi}}{\partial r} &= \vec{0} & \frac{\partial \hat{\phi}}{\partial \theta} &= \vec{0} & \frac{\partial \hat{\phi}}{\partial \phi} &= -\hat{r} \sin \theta - \hat{\theta} \cos \theta \end{aligned} \quad (3.29)$$

Note that Symon, in his equation (3.99), is inconsistent in his use of the “zero vector” notation.

We can summarize the geometrical results in three dimensions as

$$\vec{r} = x \hat{x} + y \hat{y} + z \hat{z} = q \hat{q} + z \hat{z} = r \hat{r} \quad (3.30)$$

and

$$d\vec{r} = \hat{x} dx + \hat{y} dy + \hat{z} dz = \hat{q} dq + \underbrace{\hat{\phi} (q d\phi)}_{q(d\hat{q})} + \hat{z} dz = \hat{r} dr + \underbrace{\hat{\theta} r d\theta + \hat{\phi} r \sin \theta d\phi}_{r(d\hat{r})} \quad (3.31)$$

## 4 Conservation Theorems in Two or More Dimensions

We continue our look at the analogies between one- and three-dimensional mechanics by looking at some of the conservation theorems.

### 4.1 Force and Momentum

In one dimension, we saw that we can integrate

$$F = \frac{dp}{dt} \quad (4.1)$$

to get

$$\int_{t_1}^{t_2} F(t) dt = p(t_2) - p(t_1) \quad (4.2)$$

and the same thing carries over perfectly well into two or more dimensions; we can integrate

$$\vec{F} = \frac{d\vec{p}}{dt} \quad (4.3)$$

to get

$$\int_{t_1}^{t_2} \vec{F}(t) dt = \vec{p}(t_2) - \vec{p}(t_1) \quad (4.4)$$

since the acts of differentiating and integrating the vectors  $\vec{F}$  and  $\vec{p}$  with respect to the scalar  $t$  are well-defined.

## 4.2 Work and Kinetic Energy

Recall that in one dimension, we defined the kinetic energy as

$$T = \frac{1}{2}mv^2 ; \quad (4.5)$$

this was useful because

$$\frac{dT}{dt} = mv \frac{dv}{dt} = Fv = F \frac{dx}{dt} \quad (4.6)$$

where we used Newton's second law to replace  $ma$  with  $F$ . We could then integrate this to see

$$T(t_2) - T(t_1) = \int_{t_1}^{t_2} F \frac{dx}{dt} dt = \int_{x(t_1)}^{x(t_2)} F dx = W_{1 \rightarrow 2} \quad (4.7)$$

i.e., the increase in kinetic energy is just the work done on the particle.

When generalizing the kinetic energy to two or more dimensions, we have to note that it involves the product of the velocity with itself. We know two ways to multiply vectors by vectors, and since  $\vec{v} \times \vec{v} = \vec{0}$  is not very interesting, we opt for the dot product and define

$$T = \frac{1}{2}m\vec{v} \cdot \vec{v} = \frac{1}{2}m|\vec{v}|^2 \quad (4.8)$$

This means the kinetic energy is still one-half the mass times the speed squared. Note that while force, position, velocity, momentum and acceleration all become vectors in more than one dimension, work and energy remain scalars.

The time derivative of the kinetic energy is

$$\frac{dT}{dt} = \frac{1}{2}m \left( \frac{d\vec{v}}{dt} \cdot \vec{v} + \vec{v} \cdot \frac{d\vec{v}}{dt} \right) = m \frac{d\vec{v}}{dt} \cdot \vec{v} = \vec{F} \cdot \vec{v} \quad (4.9)$$

which makes the increase in kinetic energy between two times  $t_1$  and  $t_2$

$$T(t_2) - T(t_1) = \int_{t_1}^{t_2} \vec{F} \cdot \frac{d\vec{r}}{dt} dt = \oint_{\substack{\vec{r}(t_1) \rightarrow \vec{r}(t_2) \\ \text{along } \vec{r}(t)}} \vec{F} \cdot d\vec{r} = W_{1 \rightarrow 2} \quad (4.10)$$

where the contour integral is just what we need to define the work done on the particle as it moves along its trajectory  $\vec{r}(t)$  from time  $t_1$  to time  $t_2$ .

### 4.2.1 Complications Arising from Potential Energy

In one dimension, we were able to use an indefinite integral to define the potential energy

$$V(x) = - \int F dx \quad (4.11)$$

up to an overall additive constant. This was equivalent to the force being given by

$$F = - \frac{dV}{dx} \quad (4.12)$$

and meant that the negative of the work done on a particle was just the increase in its potential energy as it moved from position  $x_1$  to position  $x_2$ :

$$V(x_2) - V(x_1) = - \int_{x_1}^{x_2} F(x) dx \quad (4.13)$$

where the arbitrary constant in the definition of  $V(x)$  cancels out.

When we try to generalize this to three dimensions, we run into trouble. Now the work is defined by a contour integral, and we can't just define  $V(\vec{r})$  as something like

$$\text{“ } - \oint \vec{F} \cdot d\vec{r} \text{ ”} \quad (4.14)$$

because there's no such thing as an “indefinite contour integral”. The contour integral depends not only on the endpoints, but the path through space taken from one to the other. The implicit definition in terms of a derivative would be

$$\text{“ } \vec{F} = - \frac{dV}{d\vec{r}} \text{ ”} \quad (4.15)$$

which is likewise in terms of something which we never defined (a derivative with respect to a vector). We could try to fall back to the change in potential energy being related to the work done:

$$V(\vec{r}_2) - V(\vec{r}_1) = - \oint_{\vec{r}_1 \rightarrow \vec{r}_2} \vec{F} \cdot d\vec{r} \quad (4.16)$$

but in general, this will not be well-defined, since the right-hand side will depend not only on the positions  $\vec{r}_1$  and  $\vec{r}_2$ , but also on the path taken between them.

### 4.3 Conservative Forces

For some force fields, however, the right-hand side of (4.16) will be the same regardless of the path taken. We'll assume this is true and see what that tells us about the force field  $F(\vec{r})$  and how it can be derived from  $V(\vec{r})$ .

First, we consider the special case where the endpoints  $\vec{r}_1$  and  $\vec{r}_2$  differ only in their  $x$  coordinate:

$$y_1 = y_2 = y_0 \quad (4.17a)$$

$$z_1 = z_2 = z_0 \quad (4.17b)$$

so that

$$\vec{r}_1 = x_1 \hat{x} + y_0 \hat{y} + z_0 \hat{z} \quad (4.18a)$$

$$\vec{r}_2 = x_2 \hat{x} + y_0 \hat{y} + z_0 \hat{z} \quad (4.18b)$$

$$(4.18c)$$

We're assuming the work done by the force field when moving between these two points is independent of the path, so let's choose the simplest one, on which only the  $x$  coordinate is changing, and parametrize it by that  $x$  coordinate:

$$\vec{r}(s) = s \hat{x} + y_0 \hat{y} + z_0 \hat{z} \quad s : x_1 \rightarrow x_2 \quad (4.19)$$



This then means

$$\frac{d\vec{r}}{ds} = \hat{x} \quad (4.20)$$

and

$$\vec{F} \cdot \frac{d\vec{r}}{ds} = \vec{F} \cdot \hat{x} = F_x \quad (4.21)$$

The difference (4.16) in the potential energies then becomes

$$V(\vec{r}_2) - V(\vec{r}_1) = - \int_{x_1}^{x_2} F_x(\vec{r}(s)) ds = V(x_2, y_0, z_0) - V(x_1, y_0, z_0) = - \int_{x_1}^{x_2} F_x(x, y_0, z_0) dx \quad (4.22)$$

Now we know that in one dimension (4.13) is equivalent to (4.12); this is the same construction, with  $y_0$  and  $z_0$  going along for the ride, so we get the same result, with the derivative with respect to  $x$  now a partial derivative at constant  $y$  and  $z$ :

$$F_x(\vec{r}) = - \frac{\partial V}{\partial x} \quad (4.23a)$$

But there is nothing special about the  $x$  direction; we could make the same argument holding  $x$  and  $z$  constant and varying  $y$ , or holding  $x$  and  $y$  constant and varying  $z$ , which means

$$F_y(\vec{r}) = - \frac{\partial V}{\partial y} \quad (4.23b)$$

$$F_z(\vec{r}) = - \frac{\partial V}{\partial z} \quad (4.23c)$$

so if the force field  $\vec{F}(\vec{r})$  is associated with the potential  $V(\vec{r})$  according to (4.16), we can write it as

$$\vec{F} = - \left( \hat{x} \frac{\partial V}{\partial x} + \hat{y} \frac{\partial V}{\partial y} + \hat{z} \frac{\partial V}{\partial z} \right) \quad (4.24)$$

Not every force field can be written in the form (4.24); for example, the force field

$$\vec{F} = A(x\hat{y} - y\hat{x}) \quad (4.25)$$

(where  $A$  is a constant) cannot. The easiest way to see this is to recall that you get the same result no matter in which order you take a mixed partial derivative:

$$\frac{\partial^2 V}{\partial x \partial y} = \frac{\partial^2 V}{\partial y \partial x} \quad (4.26)$$

but if  $\vec{F}$  is defined by (4.24), that means

$$\frac{\partial F_y}{\partial x} = - \frac{\partial^2 V}{\partial x \partial y} = - \frac{\partial^2 V}{\partial y \partial x} = \frac{\partial F_x}{\partial y} \quad (4.27)$$

On the other hand, the force (4.25) has

$$\frac{\partial F_y}{\partial x} = \frac{\partial(Ax)}{\partial x} = A \quad (4.28a)$$

$$\frac{\partial F_x}{\partial y} = \frac{\partial(-Ay)}{\partial y} = -A \neq \frac{\partial F_y}{\partial x} \quad (4.28b)$$

Similar conditions can be worked out for the other mixed partial derivatives of  $V(x, y, z)$ , resulting in the conditions

$$\frac{\partial F_y}{\partial x} = \frac{\partial F_x}{\partial y} \tag{4.29a}$$

$$\frac{\partial F_z}{\partial y} = \frac{\partial F_y}{\partial z} \tag{4.29b}$$

$$\frac{\partial F_x}{\partial z} = \frac{\partial F_z}{\partial x} \tag{4.29c}$$

in order for  $\vec{F}(\vec{r})$  to be associated with a potential. We've shown that they are necessary [no force field can be associated with a potential unless it satisfies (4.29)]; it turns out that they are also sufficient [and force field which satisfies (4.29) can be associated with a potential by (4.24)].

Friday, March 10, 2006

## 5 Fundamentals of Vector Calculus

### 5.1 The Gradient

The expression (4.24) for generating a vector field  $\vec{F}(\vec{r})$  given a scalar field  $V(\vec{r})$  allows us to give a well defined meaning to the “differentiation by a vector” we were trying to define in (4.15). It is a useful enough operation that it has its own name, the *gradient*

$$\vec{\nabla}V = \hat{x}\frac{\partial V}{\partial x} + \hat{y}\frac{\partial V}{\partial y} + \hat{z}\frac{\partial V}{\partial z} \quad (5.1)$$

The symbol  $\vec{\nabla}$  (pronounced “del”) can be thought of as a *vector differential operator*

$$\vec{\nabla} = \hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z} \quad (5.2)$$

We could also express this operator in terms of non-Cartesian basis vectors in a brute force manner as follows:

1. Invert the relationship

$$\begin{pmatrix} \hat{r} \\ \hat{\theta} \\ \hat{\phi} \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \end{pmatrix} \begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{pmatrix} \quad (5.3)$$

to get  $\hat{x}$ ,  $\hat{y}$ , and  $\hat{z}$  in terms of the spherical coordinate basis vectors.

2. Calculate the partial derivatives of the spherical coordinates with respect to the Cartesian ones and construct

$$\frac{\partial}{\partial x} = \frac{\partial r}{\partial x}\frac{\partial}{\partial r} + \frac{\partial\theta}{\partial x}\frac{\partial}{\partial\theta} + \frac{\partial\phi}{\partial x}\frac{\partial}{\partial\phi} \quad (5.4a)$$

$$\frac{\partial}{\partial y} = \frac{\partial r}{\partial y}\frac{\partial}{\partial r} + \frac{\partial\theta}{\partial y}\frac{\partial}{\partial\theta} + \frac{\partial\phi}{\partial y}\frac{\partial}{\partial\phi} \quad (5.4b)$$

$$\frac{\partial}{\partial z} = \frac{\partial r}{\partial z}\frac{\partial}{\partial r} + \frac{\partial\theta}{\partial z}\frac{\partial}{\partial\theta} + \frac{\partial\phi}{\partial z}\frac{\partial}{\partial\phi} \quad (5.4c)$$

3. Substitute, simplify, and rearrange.

Fortunately, there’s a much easier, clever way to do things.

Consider the the infinitesimal change in the value of the scalar field  $V(\vec{r})$  induced by an infinitesimal change

$$d\vec{r} = \hat{x}dx + \hat{y}dy + \hat{z}dz \quad (5.5)$$

position vector  $\vec{r}$ . We know from the chain rule of multi-variable calculus that

$$dV = \frac{\partial V}{\partial x}dx + \frac{\partial V}{\partial y}dy + \frac{\partial V}{\partial z}dz = (\vec{\nabla}V) \cdot (d\vec{r}) \quad (5.6)$$

But we also know that in spherical coordinates,

$$d\vec{r} = \hat{r} dr + \hat{\theta} r d\theta + \hat{\phi} r \cos \theta d\phi \quad (5.7)$$

which means

$$dV = (\vec{\nabla}V) \cdot (d\vec{r}) = (\hat{r} \cdot \vec{\nabla}V)dr + (\hat{\theta} \cdot \vec{\nabla}V)r d\theta + (\hat{\phi} \cdot \vec{\nabla}V)r \sin \theta d\phi \quad (5.8)$$

On the other hand, the chain rule tells us

$$dV = \frac{\partial V}{\partial r} dr + \frac{\partial V}{\partial \theta} d\theta + \frac{\partial V}{\partial \phi} d\phi \quad (5.9)$$

Since the spherical coordinate basis is orthonormal, this gives us the form of the gradient in spherical coordinates as

$$\vec{\nabla}V = \hat{r} \frac{\partial V}{\partial r} + \hat{\theta} \frac{1}{r} \frac{\partial V}{\partial \theta} + \hat{\phi} \frac{1}{r \sin \theta} \frac{\partial V}{\partial \phi} \quad (5.10)$$

## 5.2 Divergence, Curl, and All That

Thinking of  $\vec{\nabla}$  as a vector operator, we can also define how it acts on vector fields by analogy with the dot and cross products:

$$\vec{\nabla} \cdot \vec{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \quad (5.11)$$

and

$$\vec{\nabla} \times \vec{A} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix} = (\partial_y A_z - \partial_z A_y)\hat{x} + (\partial_z A_x - \partial_x A_z)\hat{y} + (\partial_x A_y - \partial_y A_x)\hat{z} \quad (5.12)$$

where we have defined the phenomenally useful notation

$$\partial_x = \frac{\partial}{\partial x} \quad (5.13)$$

There's a whole rich subject of vector calculus, and the relationship between derivatives and integrals, for which I refer you to section 3.6 of Symon, your Calc III text, and lecture notes from Fall 2002.

The one thing that's of interest to us right now is the observation that the conditions (4.29) for a force field  $\vec{F}(\vec{r})$  to be conservative are just equivalent to

$$\vec{\nabla} \times \vec{F} = \vec{0} \quad (5.14)$$

This is associated with the vector calculus identity

$$\vec{\nabla} \times \vec{\nabla}V = \vec{0} \quad (5.15)$$

which is usually stated as a consequence of the fact that a vector crossed with itself is zero  $\vec{A} \times \vec{A}$ , but note that we have to be a little careful because the vector in question here is an operator  $\vec{\nabla}$ . For instance, if the operator were  $r\vec{\nabla}$ , the result would be

$$r\vec{\nabla} \times (r\vec{\nabla}V) = r(\vec{\nabla}r) \times (\vec{\nabla}V) \neq \vec{0} \quad (5.16)$$

The real reason for this identity is once again the fact that partial derivatives “commute”, e.g., (4.26).

### 5.3 Practical Calculation of Potential Energy

We end our consideration of conservative forces with an illustration of a technique to determine the potential from which a conservative force field was generated. One method, by integrating from the origin along a conveniently chosen path to an arbitrary point, is illustrated in Section 3.12 of Symon. We illustrate here an alternative method, using the same force.

The force field in the example is

$$\vec{F} = ay(y^2 - 3z^2)\hat{x} + 3ax(y^2 - z^2)\hat{y} - 6axyz\hat{z} \quad (5.17)$$

which means the potential energy  $V(x, y, z)$  has the derivatives

$$\frac{\partial V}{\partial x} = ay(3z^2 - y^2) \quad (5.18a)$$

$$\frac{\partial V}{\partial y} = 3ax(z^2 - y^2) \quad (5.18b)$$

$$\frac{\partial V}{\partial z} = 6axyz \quad (5.18c)$$

If we knew the derivative of a one-dimensional function, we could take the indefinite integral and obtain the function up to an additive integration constant. We do basically the same thing here, integrating with respect to one variable at a time and holding the other ones constant.

First, we treat  $y$  and  $z$  as constants and integrate the  $x$  derivative:

$$V(x, y, z) = \int \frac{\partial V(x, y, z)}{\partial x} dx = \int ay(3z^2 - y^2) dx = axy(3z^2 - y^2) + A(y, z) \quad (5.19)$$

When we do the indefinite integral we get an arbitrary integration constant, but since  $y$  and  $z$  are constants for the purposes of the partial derivative, the integration constant can in principle depend on  $y$  and  $z$ .

We can use (5.18b) to determine a condition the  $A(y, z)$  introduced in (5.19):

$$3ax(z^2 - y^2) = \frac{\partial V}{\partial y} = 3axz^2 - 3xy^2 + \frac{\partial A}{\partial y} \quad (5.20)$$

we solve this to find

$$\frac{\partial A}{\partial y} = 0 \quad (5.21)$$

Now, ordinarily we'd integrate this, and get an integration "constant" which depends only on  $z$ , but here life is easier because the partial derivative vanishes, so there is no  $y$  dependence in  $A(y, z)$ :

$$A(y, z) = B(z) \quad (5.22)$$

and thus

$$V(x, y, z) = axy(3z^2 - y^2) + B(z) \quad (5.23)$$

Now we do the same thing as before, applying (5.18c) to (5.23) to find

$$6axyz = \frac{\partial V}{\partial z} = 6axyz + \frac{dB}{dz} \quad (5.24)$$

or

$$\frac{dB}{dz} = 0 \quad (5.25)$$

Again, we'd normally have to do an integral here, but this time, it's just telling us that  $B(z)$  is a constant

$$B(z) = C \quad (5.26)$$

i.e.,

$$V(x, y, z) = axy(3z^2 - y^2) + C \quad (5.27)$$

Note that in this case, the constant is just the value of the potential at the origin:

$$V(0, 0, 0) = C \quad (5.28)$$

so that

$$V(x, y, z) = V(0, 0, 0) + axy(3z^2 - y^2) \quad (5.29)$$

Typically in a physical problem, one chooses the constant in a potential so that the potential is zero at some convenient point, usually the origin or "infinity".

The same principle applies in non-Cartesian coordinates, except that you have to be careful about the definition of the gradient. So for example, in spherical coordinates,

$$\vec{F} = F_r \hat{r} + F_\theta \hat{\theta} + F_\phi \hat{\phi} = -\vec{\nabla}V = -\left(\frac{\partial}{\partial r} + \frac{1}{r} \frac{\partial}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi}\right) V(r, \theta, \phi) \quad (5.30)$$

so

$$\frac{\partial V}{\partial r} = -F_r(r, \theta, \phi) \quad (5.31a)$$

$$\frac{\partial V}{\partial \theta} = -rF_\theta(r, \theta, \phi) \quad (5.31b)$$

$$\frac{\partial V}{\partial \phi} = -r \sin \theta F_\phi(r, \theta, \phi) \quad (5.31c)$$

and we can then use the same multi-variable calculus technique to find  $V(r, \theta, \phi)$  up to an additive constant, given its partial derivatives.

A good thing to keep in mind is that vector calculus treats Cartesian and non-Cartesian coordinate systems differently, but multi-variable calculus does not.

## 5.4 Application of Potential Energy

If we are dealing with a conservative force field, one which is associated with a potential energy:

$$\vec{F} = -\vec{\nabla}V \quad (5.32)$$

then conservation of energy holds just as in one dimension:

$$T_2 - T_1 = W_{1 \rightarrow 2} = \int_{1 \rightarrow 2} \vec{F} \cdot d\vec{r} = -[V(\vec{r}_2) - V(\vec{r}_1)] \quad (5.33)$$

and therefore

$$T_1 + V_1 = T_2 + V_2 \quad (5.34)$$

At this point, we've more or less covered Symon sections 3.1-3.7 (except for the part about torque in section 3.3), 3.9, and 3.12. Sections 3.10 and 3.11 illustrate some simple and non-simple applications of the methods (the not-so-simple one also appearing on a homework), and you should read about these on your own. We will now consider rotational motion (Symon sections 3.3 and 3.8) which will prepare us for a major sub-topic, central force motion, which is described in sections 3.13 through 3.16.

Monday, March 13, 2006

## 6 Rotational Motion

In introductory mechanics, one usually learns about rotation around an axis or rotation in a plane. In general, rotational motion can be described with vector quantities which deal with rotations about all possible axes through a given origin. Our goal here is to bridge the gap between these two descriptions.

### 6.1 Torque

There are a few different ways to define torque about an axis, but a convenient one for our purposes is

$$(\text{Torque about axis}) = (\text{Distance from axis})(\text{Perpendicular component of force}) \quad (6.1)$$

where “perpendicular” means “perpendicular to the direction from the point of application to the closest point on the axis” and also “perpendicular to the axis”. We’ll eventually want a vector expression for this, but let’s start by looking at it in a convenient coordinate system. The natural one to use is cylindrical coordinates with the  $z$  axis along the axis of rotation. For a given point, the distance to the axis is the  $q$  coordinate of that point. The “perpendicular” direction at that point is the  $\phi$  direction, so the perpendicular component of the force is just  $F_\phi$ . This means that if we call the torque about the  $z$  axis  $N_z$  (using  $N$  to stand for torque for reasons which completely escape me),

$$N_z = q F_\phi \quad (6.2)$$

Notice that this also has the sign we’d like; if  $F_\phi > 0$ , we are applying positive torque around the  $z$  axis as defined by the right-hand rule.

Now, if we think for a minute about the fact that  $q$  and  $F_\phi$  are components of vectors

$$\vec{F} = F_q \hat{q} + F_\phi \hat{\phi} + F_z \hat{z} \quad (6.3a)$$

$$\vec{r} = q \hat{q} + z \hat{z} \quad (6.3b)$$

we’ll realize that the combination  $q F_\phi$  is just the  $z$  component of the cross product  $\vec{r} \times \vec{F}$ :

$$\hat{z} \cdot (\vec{r} \times \vec{F}) = q F_\phi \quad (6.4)$$

[Ordinarily there would be another term because  $\hat{z} \cdot (\vec{A} \times \vec{B}) = A_q B_\phi - A_\phi B_q$ , but the position vector has no  $\phi$  component:  $\hat{\phi} \cdot \vec{r} = 0$ .]

So this means the torque about the  $z$  axis is

$$N_z = \hat{z} \cdot (\vec{r} \times \vec{F}) \quad (6.5)$$

But now the only reference to the coordinate system is the unit vector  $\hat{z}$  along the  $z$  axis (and the origin implicit in the position vector  $\vec{r}$ ). So in general we can say that the torque  $N_n$  around some axis through the origin parallel to a constant unit vector  $\hat{n}$  is

$$N_n = \hat{n} \cdot (\vec{r} \times \vec{F}) \quad (6.6)$$



Of course, the notation is supposed to be suggestive, and we can encode the torque about all possible axes through the origin if we define the vector

$$\vec{N} = \vec{r} \times \vec{F} \quad (6.7)$$

which is the torque.

## 6.2 Angular Momentum

The geometry of the definition of angular momentum is the same. The angular momentum about an axis is the distance from that axis times the perpendicular component of the momentum, so working once more in cylindrical coordinates

$$L_z = q p_\phi \quad (6.8)$$

Now, since we know that the velocity in cylindrical coordinates is

$$\vec{v} = \dot{q}\hat{q} + q\dot{\phi}\hat{\phi} + \dot{z}\hat{z} \quad (6.9)$$

we can write

$$p_\phi = mv_\phi = mq\dot{\phi} \quad (6.10)$$

to get the perhaps more familiar

$$L_z = mq^2\dot{\phi} \quad (6.11)$$

By the same vector argument as we used for torque, we can see that

$$L_z = \hat{z} \cdot (\vec{r} \times \vec{p}) \quad (6.12)$$

and define a vector angular momentum

$$\vec{L} = \vec{r} \times \vec{p} \quad (6.13)$$

## 6.3 Relationship of Angular Momentum and Torque

We can use Newton's second law to show that angular momentum and torque are related in the same way as linear momentum and force.

$$\frac{d\vec{L}}{dt} = \frac{d}{dt} (\vec{r} \times \vec{p}) = \frac{d\vec{r}}{dt} \times \vec{p} + \vec{r} \times \frac{d\vec{p}}{dt} \quad (6.14)$$

Now, the first term is

$$\frac{d\vec{r}}{dt} \times \vec{p} = \vec{v} \times (m\vec{v}) = m\vec{v} \times \vec{v} = \vec{0} \quad (6.15)$$

while we can use Newton's second law to make the replacement

$$\frac{d\vec{p}}{dt} = \vec{F} \quad (6.16)$$

and thus obtain

$$\frac{d\vec{L}}{dt} = \vec{r} \times \vec{F} = \vec{N} \quad (6.17)$$

An important special case of this is when the torque vanishes, i.e.,  $\vec{r} \times \vec{F} = \vec{0}$ . Then the vector angular momentum  $\vec{L}$  is a constant of the motion.

## 7 Central Force Motion

An important physical problem is that of a particle moving in an fixed force field, where the magnitude of the force depends only on the distance from a specified point, and the direction of the force is always either towards or away from that point. Mathematically, if we define a spherical coordinate system with the origin at the point in question, this means

$$\vec{F}(\vec{r}) = F(r)\hat{r} \quad (7.1)$$

Note that  $F(r)$  is the  $r$ -component of the force  $\vec{F}$  and is *not* the magnitude. In particular, it can be positive or negative.

The most obvious physical example is that of a planet orbiting a star. It's a decent approximation to assume that the star doesn't move much, and therefore the planet is moving in the fixed gravitational field of the star:

$$\vec{F} = -\frac{GMm}{r^2}\hat{r} \quad (7.2)$$

We'll see in chapter four that even if  $m \ll M$  we can still describe the two-body problem exactly in terms of an equivalent one-body problem with a particle orbiting a fixed center of force.

Thursday, March 16, 2006

## 7.1 Conservation of Energy

It's easy to show that (7.1) is a conservative force. We could calculate the curl  $\vec{\nabla} \times \vec{F}$ , but it's actually more straightforward just to construct the potential explicitly. Since the force is determined by one function of one variable, we can take its indefinite integral with respect to  $r$ :

$$V(r) = - \int F(r) dr \quad (7.3)$$

We can show explicitly that this is indeed the potential energy by taking its gradient:

$$-\vec{\nabla}V = -\frac{\partial V}{\partial r}\hat{r} = F(r)\hat{r} = \vec{F} \quad (7.4)$$

This means that the total energy

$$E = T + V = \frac{1}{2}m\vec{v} \cdot \vec{v} + V(r) \quad (7.5)$$

is a constant of the motion.

## 7.2 Conservation of Angular Momentum

We can also show that the torque (about the origin) associated with this force is zero:

$$\vec{N} = \vec{r} \times \vec{F} = (r\hat{r}) \times (F(r)\hat{r}) = rF(r)(\hat{r} \times \hat{r}) = \vec{0} \quad (7.6)$$

This means that the vector angular momentum

$$\vec{L} = \vec{r} \times m\vec{v} \quad (7.7)$$

is a constant. As a consequence (which we will show):

1. The motion is confined to a plane.
2. There is an additional non-vector constant of the motion

To see that the motion remains confined to a plane, align the coordinate axes so that initially (at  $t = 0$ ), both the position and the velocity lie in the  $xy$  plane:

$$z(0) = \hat{z} \cdot \vec{r}(0) = 0 \quad (7.8a)$$

$$v_z(0) = \hat{z} \cdot \vec{v}(0) = 0 \quad (7.8b)$$

Now define a cylindrical coordinate system and calculate the components of the angular momentum (at any time) in that system:

$$\vec{L} = m\vec{r} \times \vec{v} = m \begin{vmatrix} \hat{q} & \hat{\phi} & \hat{z} \\ q & 0 & z \\ v_q & v_\phi & v_z \end{vmatrix} = -mzv_\phi\hat{q} + m(zv_q - qv_z)\hat{\phi} + mqv_\phi\hat{z} \quad (7.9)$$

We evaluate this at the initial time and see

$$\vec{L}(0) = mq(0)v_\phi(0)\hat{z} \quad (7.10)$$

where we have used (7.8a) to show that the other components of the angular momentum vanish. But  $\vec{L}(0)$  is a constant of the motion, which means

$$\vec{L}(t) = \vec{L}(0) = mq(0)v_\phi(0)\hat{z} \quad (7.11)$$

Because we're working in cylindrical coördinates, we need to be a little careful about just setting the components of  $\vec{L}(0)$  and  $\vec{L}(t)$  equal to each other, because the basis vectors  $\hat{q}$  and  $\hat{\phi}$  depend on the position, which in turn depends on time. But fortunately the only basis vector appearing in (7.11) is  $\hat{z}$ , which is a constant, and  $\hat{q}$  and  $\hat{\phi}$ , while they will depend on  $t$ , are always perpendicular to  $\hat{z}$ . This means we can say that  $\vec{L}(t)$  has no  $q$  or  $\phi$  component, and that its  $z$  component equals that of

$$-mz(t)v_\phi(t) = 0 \quad (7.12a)$$

$$mz(t)v_q(t) - mq(t)v_z(t) = 0 \quad (7.12b)$$

$$mq(t)v_\phi(t) = mq(0)v_\phi(0) \quad (7.12c)$$

Now, unless the initial angular momentum vanishes (which would be a special case, where the particle was initially moving straight towards or away from the origin), the product  $qv_\phi$ , which is a constant, would remain non-zero, which means neither  $q$  nor  $v_\phi$  can be zero. This means that (7.12a) tells us  $z(t) = 0$  for all time, which in turn tells us, along with (7.12b), that  $v_z(t) = 0$  for all time. In other words, if the motion starts out in the  $xy$  plane and the angular momentum is a constant (which it is for a central force) the motion remains in the  $xy$  plane forever.

We needed two of the components of the conserved angular momentum to tell us this; the third tells us that

$$L_z = mqv_\phi = \text{constant} \quad (7.13)$$

As a consequence of the fact that  $z \equiv 0$  in this coördinate system,

$$r = \sqrt{q^2 + z^2} \equiv q \quad (7.14)$$

and this is basically a two-dimensional problem described in plane polar coördinates  $(r, \phi)$ .

### 7.3 The Nature of the Differential Equations

In plane polar coördinates, thanks to the time derivatives

$$\frac{d\hat{r}}{dt} = \dot{\phi}\hat{\phi} \quad (7.15a)$$

$$\frac{d\hat{\phi}}{dt} = -\dot{\phi}\hat{r} \quad (7.15b)$$

of the basis vectors, we know the position, velocity and acceleration are

$$\vec{r} = r\hat{r} \quad (7.16a)$$

$$\vec{v} = \dot{r}\hat{r} + r\dot{\phi}\hat{\phi} \quad (7.16b)$$

$$\vec{a} = (\ddot{r} - r\dot{\phi}^2)\hat{r} + (r\ddot{\phi} - 2\dot{r}\dot{\phi})\hat{\phi} \quad (7.16c)$$

and so using Newton's second law

$$\vec{F} = m\vec{a} \quad (7.17)$$

we could derive the equations of motion

$$m(\ddot{r} - r\dot{\phi}^2) = F(r) \quad (7.18a)$$

$$m(r\ddot{\phi} - 2\dot{r}\dot{\phi}) = 0 \quad (7.18b)$$

These are two second-order ODEs for the two functions  $r(t)$  and  $\phi(t)$ , so we could in principle solve them subject to four initial conditions, e.g., specified values of  $r(0)$ ,  $\dot{r}(0)$ ,  $\phi(0)$  and  $\dot{\phi}(0)$ .

However, conservation of energy and angular momentum allows us to cast the problem more simply. A given trajectory has a given value of  $E$  and  $L$ , which are related to  $r(t)$  and  $\phi(t)$  by

$$E = \frac{1}{2}m\vec{v} \cdot \vec{v} + V(r) = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\phi}^2 + V(r) \quad (7.19a)$$

$$L = mrv_{\phi} = mr^2\dot{\phi} \quad (7.19b)$$

These are then, for given values of  $E$  and  $L$ , two *first-order* ODES for  $r(t)$  and  $\phi(t)$ . Given  $E$  and  $L$  plus two initial conditions such as values of  $r(0)$ ,  $\dot{r}(0)$ , and  $\phi(0)$ , we can solve them for  $r(t)$  and  $\phi(t)$ .

The explicit formal solution is as follows:

First, we note that since (7.19b) contains only  $r$  and  $\dot{\phi}$ , we can use it to solve for  $\dot{\phi}$  and then remove it from (7.19a):

$$\dot{\phi} = \frac{L}{mr^2} \quad (7.20)$$

and therefore

$$E = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2 \left( \frac{L}{mr^2} \right)^2 + V(r) = \frac{1}{2}m\dot{r}^2 + \frac{L^2}{2mr^2} + V(r) \quad (7.21)$$

Equation (7.21) is then a single first-order ODE for  $r(t)$ . We can solve for

$$\dot{r} = \pm \sqrt{\frac{2}{m}} \sqrt{E - \left[ V(r) + \frac{L^2}{2mr^2} \right]} = \frac{dr}{dt} \quad (7.22)$$

which can then be solved by integration

$$\int_{r(0)}^{r(t)} \frac{dr}{\sqrt{E - \left[ V(r) + \frac{L^2}{2mr^2} \right]}} = \int_0^t \sqrt{\frac{2}{m}} dt' = t \sqrt{\frac{2}{m}} \quad (7.23)$$

Once we've got  $r(t)$ , we can then plug it into (7.20) and integrate to get

$$\phi(t) = \phi(0) + \int_0^t \dot{\phi}(t') dt' = \phi(0) + \int_0^t \frac{L}{mr^2(t')} dt' \quad (7.24)$$

## 7.4 Effective Potential

We've just described formally how one could solve a central force motion problem. It may often turn out, though, that the integrals (7.23) and (7.24) can't be evaluated in closed form. It is thus useful to be able to gain more insight into the problem by other methods of analysis.

Look at the conserved total energy (7.21). Although we started out with a three-dimensional problem and thus among other things a different expression for the kinetic energy, the end product actually looks a lot like the energy in a one-dimensional problem. Recall that this is

$$E = \frac{1}{2}m\dot{x}^2 + V(x) \quad (7.25)$$

while our total energy can be written in the form

$$E = \frac{1}{2}m\dot{r}^2 + V_{\text{eff}}(r) \quad (7.26)$$

where

$$V_{\text{eff}}(r) = V(r) + \frac{L^2}{2mr^2} \quad (7.27)$$

is the so-called “effective potential”. We've reduced our three-dimensional problem to an equivalent one-dimensional problem where the only complication is the addition to the potential energy of a term, dependent on the angular momentum  $L$ , which accounts for the rotational motion. This is often referred to as the “centrifugal barrier”, since it blows up at small  $r$ .

Equation (7.26) allows us to use all the same methods as we did back in one-dimensional potential problems, but with  $V_{\text{eff}}$  as the potential. For example, at a given energy  $E$ , the turning points  $r_t$  are the solutions to  $V_{\text{eff}}(r_t) = E$ , and a stable equilibrium occurs at a minimum of  $V_{\text{eff}}$ . Now, of course, equilibrium means constant  $r$ . The angular coordinate  $\phi$ , meanwhile, is always changing at a rate

$$\dot{\phi} = \frac{L}{mr^2} \quad (7.28)$$

So the “equilibrium” corresponds to a circular orbit, around which the particle moves at a constant angular speed. In general, for a given value of  $L$ , which fixes a given  $V_{\text{eff}}(r)$ , this can only occur at particular values of  $r$ .

As in the one-dimensional case, if we are near the equilibrium radius  $r_e$  for a given  $V_{\text{eff}}(r)$ , the  $r$  value of the particle will oscillate about  $r_e$  with a frequency

$$\omega_R \approx \sqrt{\frac{V_{\text{eff}}''(r_e)}{m}} \quad (7.29)$$

On the other hand, the angular frequency of the azimuthal motion will be

$$\omega_\Phi \approx \frac{L}{mr_e^2} \quad (7.30)$$

and in general these two will not be the same, so the orbit will not close, leading to something like Symon's figure 3.34.

Monday, March 20, 2006

## 7.5 Finding the Trajectory

As mentioned above, the integrals (7.23) and (7.24) for  $r(t)$  and  $\phi(t)$  often can't be evaluated in closed form. Sometimes it's easier to find the shape of the orbit  $r(\phi)$  and not worry about when the particle reaches which position. For given  $E$  and  $L$ , we get a first-order equation

$$\frac{dr}{d\phi} = \frac{\dot{r}}{\dot{\phi}} = \frac{\pm\sqrt{2/m}\sqrt{E - V(r) - L^2/2mr^2}}{L/2mr^2} \quad (7.31)$$

In practice, it often turns out to be even easier to work in terms of  $u(\phi) = 1/r(\phi)$ . This is because

$$\frac{du}{d\phi} = -\frac{1}{r^2} \frac{dr}{d\phi} = -\frac{\dot{r}}{r^2\dot{\phi}} = \mp \frac{\sqrt{2[E - V(r)]/m - (L/mr)^2}}{L/m} = \mp \sqrt{\frac{2m[E - V(1/u)]}{L^2} - u^2} \quad (7.32)$$

We'll see this in detail when we consider the potential energy which is meaningful for gravity.

## 7.6 Kepler's Second Law

In the 17th century, Johannes Kepler discovered three empirical properties of planetary orbits, which are known as Kepler's Laws of Planetary Motion. They are:

1. Planetary orbits are in the shape of ellipses with the Sun at one focus.
2. A planet moves around its orbit at such a rate that the line from the planet to the Sun sweeps out equal areas in equal intervals of time.
3. The cube of the semimajor axis of a planet's orbit is proportional to the square of its orbital period.

Isaac Newton subsequently used these laws to deduce the nature of the gravitational interaction. The first and third laws turn out to be consequences of the particular spatial dependence of the gravitational interaction, but the second holds for any central force, being a consequence of the conservation of angular momentum.

This can be seen by considering the wedge swept out in a time  $dt$ . For sufficiently small time, the change in radius is negligible, and the area is approximately that of a circular wedge of radius  $r$  and angular extent  $d\phi = \dot{\phi} dt$ . This is approximately a triangle height  $r$  and base  $r d\phi$ , which has area

$$dA = \frac{1}{2}r(r d\phi) = \frac{1}{2}r^2\dot{\phi} dt = \frac{L}{2m} dt \quad (7.33)$$

and since  $L$  is a constant for a central force, so is

$$\frac{dA}{dt} = \frac{L}{2m} \quad (7.34)$$

Friday, March 24, 2006

## 8 Inverse-Square-Law Forces

A very useful special case of a central force is one inversely proportional to the square of the distance:

$$F(r) = -\frac{GMm}{r^2} \quad (8.1)$$

This is the gravitational force on a particle of mass  $m$  due to a point source of mass  $M$  fixed at the origin.

The potential energy corresponding to this force is

$$V(r) = \int \frac{GMm}{r^2} dr = -\frac{GMm}{r} + \text{const} \quad (8.2)$$

The conventional choice of the zero of the potential sets the constant to zero:

$$V(r) = -\frac{GMm}{r} \quad (8.3)$$

this has the property that  $V(\infty) = 0$ .

### 8.1 Finding the Turning Points

We can examine the behavior by looking at the effective potential

$$V_{\text{eff}}(r) = -\frac{GMm}{r} - \frac{L^2}{2mr^2} \quad (8.4)$$

At small  $r$ , as long as  $L \neq 0$ , the centrifugal term blows up faster and dominates:

$$V_{\text{eff}}(r) \xrightarrow{r \rightarrow 0} -\frac{L^2}{2mr^2} \quad (8.5)$$

At large  $r$ , the other term goes to zero more slowly, and eventually dominates:

$$V_{\text{eff}}(r) \xrightarrow{r \rightarrow \infty} -\frac{GMm}{r} \quad (8.6)$$

Qualitatively, then we see the behavior that there is a local minimum  $r_e$  for which  $V(r_e) < 0$ .

Orbits with  $E > 0$  have one inner turning point  $r_t = r_{\min}$  at which  $V_{\text{eff}}(r_t) = E$ . (The value of  $r_t$  will of course depend on  $E$  and  $L$ .)

Orbits with  $E < 0$  have two turning points  $r_{\min} < r_e < r_{\max}$  at which  $V_{\text{eff}}(r_t) = E$ .

We write the equation for the turning points as

$$E = -\frac{GMm}{r_t} + \frac{L^2}{2mr_t^2} \quad (8.7)$$

Again, it turns out to be easier to work in terms of  $u = 1/r$ , in which case the turning points are given by

$$\frac{L^2}{2m}u_t^2 - GMmu_t - E = 0 \quad (8.8)$$



or

$$u_t^2 - \frac{2GMm^2}{L^2}u_t - \frac{2mE}{L^2} = 0 \quad (8.9)$$

The solutions to this quadratic equation are

$$u_t = \frac{GMm^2}{L^2} \pm \sqrt{\left(\frac{GMm^2}{L^2}\right)^2 + \frac{2mE}{L^2}} = \frac{GMm^2}{L^2} \left(1 \pm \sqrt{1 + \frac{2EL^2}{(GMm)^2m}}\right) \quad (8.10)$$

Note that if

$$E < -\frac{(GMm)^2m}{2L^2} \quad (8.11)$$

both roots are imaginary, which corresponds to an energy below the minimum of the effective potential.

As an exercise, check that the dimensions work out.

Define the positive constants

$$\alpha = \frac{L^2}{GMm^2} \quad (8.12a)$$

$$\varepsilon = \sqrt{1 + \frac{2EL^2}{(GMm)^2m}} \quad (8.12b)$$

so that

$$u_t = \alpha^{-1}(1 \pm \varepsilon) \quad (8.13)$$

Now, note that

- If  $E < 0$ ,  $\varepsilon < 1$  and both roots are positive, meaning there are two turning points, as we expect:

- The maximum value of  $u$ , which corresponds to the minimum value of  $r$ , the closest approach of the particle to the origin, which is known as *periapse*, is

$$u_p = \alpha^{-1}(1 + \varepsilon) \quad (8.14)$$

- The minimum value of  $u$ , which corresponds to the maximum value of  $r$ , the farthest displacement of the particle from the origin, which is known as *apoapse*, is

$$u_a = \alpha^{-1}(1 - \varepsilon) \quad (8.15)$$

- If  $E = 0$ ,  $\varepsilon = 1$  and one root is zero, meaning that the minimum value of  $u$  is zero and the maximum value of  $r$  is infinite, so there is no apoapse, and the periapse is still

$$u_p = \alpha^{-1}(1 + \varepsilon) \quad (8.16)$$

- If  $E > 0$ ,  $\varepsilon > 1$  and one root is negative, meaning that the minimum *allowed* value of  $u$  is still zero and the maximum allowed value of  $r$  is still infinite, so as in the marginal  $E = 0$  case, there is no apoapse and the periapse is located at

$$u_p = \alpha^{-1}(1 + \varepsilon) \quad (8.17)$$

## 8.2 Finding the Trajectory

Now we'd like to do the integral to find  $u(\phi)$  and thus  $r(\phi)$ . In all three cases,  $\frac{du}{d\phi}$  vanishes at the periapse  $u_p = \alpha^{-1}(1 + \varepsilon)$ , so let's use that as our lower limit of integration. Let's take the upper limit of integration to be some  $\phi > \phi_p$  between the two turning points, so that  $\frac{du}{d\phi} \leq 0$  (since  $u_p$  is the maximum value of  $u$ ) and

$$\phi - \phi_p = \int_{\phi_p}^{\phi} d\phi' = \int_{u_p}^{u(\phi)} \frac{du}{du/d\phi} = - \int_{(1+\varepsilon)/\alpha}^{u(\phi)} \frac{du}{\sqrt{\frac{2m(E+GMmu)}{L^2} - u^2}} \quad (8.18)$$

where we have used  $V(r) = -\frac{GMm}{r} = -GMmu$ .

Now, the integral will look a lot simpler if we substitute for  $E$  and  $L$  in terms of  $\alpha^{-1}$  and  $\varepsilon$ . We can do the algebra using the definitions (8.12), but in fact we should be able to see that the argument in the square root must be

$$\frac{2m(E + GMmu)}{L^2} - u^2 = (u_+ - u)(u - u_-) \quad (8.19)$$

where

$$u_{\pm} = \alpha^{-1}(1 \pm \varepsilon) \quad (8.20)$$

This is because we got  $u_{\pm}$  in the first place by setting the quadratic expression appearing inside the square root to zero, so it has to vanish if  $u = u_+$  or  $u = u_-$ . After that, it's just a question of getting the right coefficient (i.e.,  $-1$ ) for the  $u^2$  term. (If you don't believe me, you can do the algebra and check.)

Anyway,

$$(u_+ - u)(u - u_-) = -u^2 + (u_+ + u_-)u - u_+u_- = -u^2 + 2\alpha^{-1}u - \alpha^{-2}(1 - \varepsilon^2) = \alpha^{-2}\varepsilon^2 - (\alpha^{-1} - u)^2 \quad (8.21)$$

so

$$\phi - \phi_p = - \int_{\alpha^{-1}(1+\varepsilon)}^{u(\phi)} \frac{du}{\sqrt{\alpha^{-2}\varepsilon^2 - (\alpha^{-1} - u)^2}} \quad (8.22)$$

This is a doäble integral; making the identification  $x = \alpha^{-1} - u$ , it's of the form

$$\int \frac{dx}{a^2 - x^2} \quad (8.23)$$

which calls for the trigonometric substitution

$$x = a \sin \theta \quad (8.24)$$

so in our case, we want to define a new integration variable  $\psi$  by

$$\alpha^{-1} - u = \alpha^{-1}\varepsilon \sin \psi \quad (8.25)$$

In that case,

$$- du = \alpha^{-1}\varepsilon \cos \psi d\psi \quad (8.26)$$

and

$$\sqrt{\alpha^{-2}\varepsilon^2 - (\alpha^{-1} - u)^2} = \alpha^{-1}\varepsilon\sqrt{1 - \sin^2\psi} = \alpha^{-1}\varepsilon\cos\psi \quad (8.27)$$

We'll see we're justified in calling this  $\cos\psi$  and not  $-\cos\psi$  when we consider the limits of integration. The lower limit  $\psi_p$  obeys

$$\sin\psi_p = \frac{\alpha^{-1} - \alpha^{-1}(1 + \varepsilon)}{\alpha^{-1}\varepsilon} = 1 \quad (8.28)$$

so if we use

$$\psi_p = \frac{\pi}{2} \leq \psi \leq -\frac{\pi}{2} \quad (8.29)$$

we're justified in assuming that  $\cos\psi \geq 0$ .

Putting it all together, the integral is

$$\phi - \phi_p = \int_{\pi/2}^{\sin^{-1}\left(\frac{\alpha^{-1}-u(\phi)}{\alpha^{-1}\varepsilon}\right)} d\psi = \sin^{-1}\left(\frac{\alpha^{-1}-u(\phi)}{\alpha^{-1}\varepsilon}\right) - \frac{\pi}{2} \quad (8.30)$$

or, to solve for  $u(\phi)$ ,

$$\frac{\alpha^{-1} - u(\phi)}{\alpha^{-1}\varepsilon} = \sin\left(\frac{\pi}{2} + \phi - \phi_p\right) = \cos(\phi - \phi_p) \quad (8.31)$$

which can be solved to give

$$u(\phi) = \frac{1 + \varepsilon \cos(\phi - \phi_p)}{\alpha} \quad (8.32)$$

or, in terms of the radial coördinate,

$$r(\phi) = \frac{\alpha}{1 + \varepsilon \cos(\phi - \phi_p)} \quad (8.33)$$

Note that while our definition made some simplifying assumptions about where in the orbit we were, the solution actually holds in general.

### 8.3 Orbits in an Inverse-Square Force Field

Equation (8.33) is the equation, in polar coördinates, of a *conic section of eccentricity*  $\varepsilon$  and *latus rectum*  $\alpha$ . Specifically

- If  $\varepsilon > 1$ , this is a hyperbola with one focus at the origin
- If  $\varepsilon = 1$ , this is a parabola with its focus at the origin
- If  $\varepsilon < 1$ , this is an ellipse with one focus at the origin

This is basically Kepler's first law. Kepler only looked at planetary orbits, so he only saw the ellipses. Long-period comets have approximately parabolic orbits, while a piece of interstellar debris on a ballistic trajectory around the Sun would have a hyperbolic orbit.

If we focus on the  $\varepsilon < 1$  case, it's useful to define the *semimajor axis*

$$a = \frac{\alpha}{1 - \varepsilon^2} \quad (8.34)$$

so that the equation of the orbit is

$$r(\phi) = \frac{a(1 - \varepsilon^2)}{1 + \varepsilon \cos(\phi - \phi_p)} \quad (8.35)$$

The distances from the origin of periapse and apoapse are

$$r_p = \frac{a(1 - \varepsilon^2)}{1 + \varepsilon} = a(1 - \varepsilon) \quad (8.36a)$$

$$r_a = \frac{a(1 - \varepsilon^2)}{1 - \varepsilon} = a(1 + \varepsilon) \quad (8.36b)$$

So that

$$r_p + r_a = 2a \quad (8.37)$$

and the semimajor axis is just half the width of the ellipse at its widest point.

## 8.4 Some notes for mechanics by Brans, March 24, 2006

As a supplement to Whelan's notes for this class, these may be helpful. There seem to be some algebraic errors in Whelan's notes starting at the top of page 41. [I think they have been fixed now—JTW]

### 8.4.1 Effective potential and “centrifugal force”

The motion of a particle in a central force necessarily conserves angular momentum,  $L$ , and the motion in the radial direction can be characterized by a total energy

$$E = \frac{1}{2}mv^2 + V_{\text{eff}}(r),$$

with

$$V_{\text{eff}}(r) = V(r) + \frac{L^2}{2mr^2}.$$

So, a co-rotating coordinate system would see the particle moving only in the  $r$  direction, one dimensional, but subject not only to the “true” force,  $F_{\text{true}} = -\frac{dV}{dr}$ , but to an additional, repulsive, force,

$$F_c = -\frac{d}{dr} \frac{L^2}{2mr^2} = \frac{L^2}{mr^3}.$$

Using  $L = mv_t r$ , (here  $v_t$  is tangential speed) this reduces to

$$F_c = mv_t^2/r,$$

which is the notorious “centrifugal force” generally vigorously eschewed as fictitious in introductory courses. It must be added, or invented, if we want to use Newton's 2nd law in a rotating reference frame. As Whelan points out in his notes, it acts as an effective barrier to getting to  $r = 0$ .

### 8.4.2 Turning points

After a motion has been reduced to one dimension, as in this central force one, we can sketch a graph of the potential ( $V_{\text{eff}}$  in this case). Recalling that kinetic energy must be non-negative, for a given constant total energy,  $E$ , the allowed/forbidden regions are those for which

$$\begin{aligned} V(r) &\leq E, & \text{allowed,} \\ V(r) &> E, & \text{forbidden.} \end{aligned}$$

The boundary points between these two are called “turning points”, because the kinetic energy, and thus the speed, must be zero at them. Using  $u = 1/r$ , the points,  $u_t$ , at which the kinetic energy vanishes satisfy

$$E = V_{\text{eff}}(1/u) = -GmMu_t + \frac{L^2}{2m}u_t^2$$

This agrees with Whelan’s (8.9), but the solutions should actually be

$$u_t = \frac{Gm^2M}{L^2} \pm \sqrt{\left(\frac{Gm^2M}{L^2}\right)^2 + \frac{2mE}{L^2}}. \quad (8.38)$$

These disagree with Whelan’s (8.10), unless we double  $G$ . I believe that this extra factor of 2 causes problems in some of the rest of his equations, including his definition of  $\alpha$ , (8.12), used later in the description of the orbit, (8.35). At any rate, (8.38) above agrees with Symon’s (3.240), using  $K = -GmM$ . [This factor of 2 error has now been corrected in this version of the notes–JTW]

### 8.4.3 Orbit geometry for Newtonian gravity.

The geometry of the orbits, i.e.,  $u(\phi)$ , as opposed to the dynamics,  $u(t)$ , are governed by equations (7.32) in Whelan or (3.233) in Symon. To see the relationship between these, square Whelan’s (7.32), using  $V(1/u) = -GMmu$ , to get

$$\left(\frac{du}{d\phi}\right)^2 + u^2 = \frac{2m}{L^2}(E + GmMu) \quad (8.39)$$

Differentiate this with respect to  $\phi$ , and divide by  $2\frac{du}{d\phi}$  to get

$$\frac{d^2u}{d\phi^2} + u = \frac{Gm^2M}{L^2} \equiv \gamma.$$

Define  $w = u - \gamma$  and get

$$\frac{d^2w}{d\phi^2} + w = 0,$$

which is the famous **harmonic oscillator equation**. Careful mathematical analysis shows that such an equation has a most general solution which involves two independent arbitrary constants. By direct substitution we find such a solution to be

$$w = u - \gamma = A \cos(\phi - \phi_p),$$

or

$$u = 1/r = \gamma + A \cos(\phi - \phi_p). \quad (8.40)$$

with  $A, \phi_p$  as yet arbitrary constants. Differentiating (8.40) with respect to  $\phi$  and using (8.39), determines

$$A^2 = \frac{2mE}{L^2} + \gamma^2 = \gamma^2 \left(1 + \frac{2mE}{L^2\gamma^2}\right),$$

which agrees with Symon, (3.241). Finally, choosing coordinates so that  $\phi_p = 0$ , the orbit can be described by

$$u = 1/r = \gamma(1 + \epsilon \cos \phi) \quad (8.41)$$

where

$$\epsilon = \sqrt{1 + \frac{2mE}{L^2\gamma^2}} \quad (8.42)$$

These last two equations describe a class of curves known as **conic sections**. The distinction between unbound ( $r = 1/u$  can go to  $\infty$ ) and bound states ( $r = 1/u < B$  for some finite  $B$ ) can be translated into  $\epsilon > 1$ , which corresponds to  $E > 0$  for unbound, and  $\epsilon \leq 1$ , which corresponds to  $E \leq 0$  for bound. Geometrically the unbound states are parabolas and hyperbolas while the bound states are ellipses.

#### 8.4.4 Some elliptical notes on ellipses.

Our equations (8.41) and (8.42) define the path of an ellipse if  $0 < \epsilon^2 < 1$ , which is true for our case. To tie this in with other conic section formulations, define

$$a = \left| \frac{GmM}{2E} \right|, b = a\sqrt{1 - \epsilon^2}.$$

You might try several exercises to understand this geometry:

- Show that these equations result in the usual formula for an ellipse,

$$\frac{\bar{x}^2}{a^2} + \frac{\bar{y}^2}{b^2} = 1, \quad (8.43)$$

if we set  $\bar{x} = x + a\epsilon$ , shifted from the focus to the geometric center, and, as usual,

$$x = r \cos \phi, \quad y = r \sin \phi.$$

- Compute the area of the ellipse. This is perhaps easiest in the  $\bar{x}, \bar{y}$  system:

$$\text{Area} = \int_a^{-a} \int_{-\sqrt{b^2 - b^2 \bar{x}^2/a^2}}^{\sqrt{b^2 - b^2 \bar{x}^2/a^2}} d\bar{y} d\bar{x}.$$

Get  $\text{Area} = \pi ab$ .

- From Kepler's equal areas law,  $d\text{Area}/dt = \frac{L}{2m}$ , a constant, so the period,  $\tau$  satisfies

$$\pi ab = \frac{L}{2m} \tau.$$

From this, and the definitions of  $a, b$  above, get Kepler's third law.

## A Appendix: Correspondence to Class Lectures

Date	Sections	Pages	Topics
2006 February 21	1–1.5	3–9	Vector Arithmetic; Bases; Dot and Cross Product
2006 February 24	1.6–2.2	10–14	Position Vector; Differentiation & Integration; Vectors in Mechanics
2006 March 3	3–3.2	15–20	Non-Cartesian Coördinates
2006 March 6	3.3–4	21–26	Spherical Coördinates; Conservation Theorems
2006 March 10	5	27–31	Fundamentals of Vector Calculus
2006 March 13	6–7	32–34	Rotational motion; Central force motion
2006 March 16	7.1–7.4	35–38	Central force motion; Effective Potential
2006 March 20	7.5–7.6	39–39	Central force trajectories; Kepler's second law
2006 March 24	8	40–46	Inverse-square-law forces