2.1 Vector-Valued Functions

Vectors differ from regular numbers because they have both a magnitude (length) and a direction. In our three dimensional world, vectors have one component for each direction, and are denoted by

$$\mathbf{r}(t) = (x, y, z)\,. \tag{1}$$

Another common notation uses the unit vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ for the x, y and z direction respectively. The vector is then written as

$$\mathbf{r}(t) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}\,. \tag{2}$$

These numbers x, y and z give a position relative to some reference point,

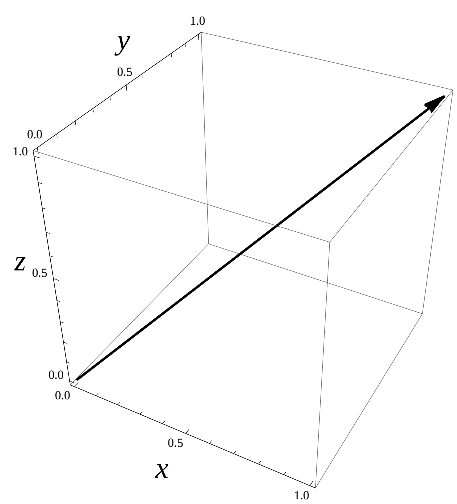

Figure 1: $\mathbf{r}(t) = \mathbf{i} + \mathbf{j} + \mathbf{k}$

which we call the origin and has coordinates $(0, 0, 0)$. An analogy is how you might give someone directions (before the advent of Google Maps!) Imagine that you lived in an apartment in a city with perpendicular streets (like American cities.) You could direct someone by saying that you live in the third street west (x position) and seventh street from the north (y position) from their apartment, and that also you live on the sixth floor (z position.) Only with all three pieces of information could your friend find you. In other words, vectors give complete information about position. Recall that a we can

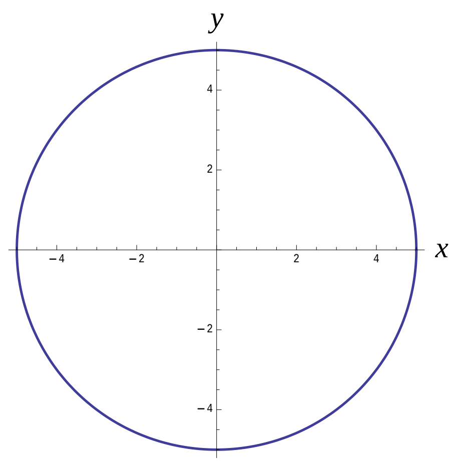

Figure 2: $\mathbf{r}(t) = 5\cos t\mathbf{i} + 5\sin t\mathbf{j}$

add two vectors $\mathbf{r}_1 = (1, 3, 1)$ and $\mathbf{r}_2 = (0, 3, -2)$ to get $\mathbf{r}_1 + \mathbf{r}_2 = (1, 6, -1)$, and so on. The length of a vector is given by its **magnitude**, the formula for which is

$$||\mathbf{r}|| = \sqrt{x^2 + y^2 + z^2}\,. \tag{3}$$

Similarly, when we think of functions, we are used to parametric functions such as

$$x(t) = 3t\,, \quad x(t) = t^2\,, \tag{4}$$

etc. However, it is reasonable to think that this can be generalised to three (or more!) dimensions. In this instance they describe shapes in three-dimensions. We define **vector-valued functions** to be functions of a real variable with several **component functions** depending on a **parametric variable** t as

$$\mathbf{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} + h(t)\mathbf{k}\,. \tag{5}$$

An important word to take note of is "real". In this course we will only be concerned with functions of real variables $t \in \mathbf{R}$. The **domain** of $\mathbf{r}(t)$, denoted $\mathcal{D}(\mathbf{r})$, is the set of values of t for which $\mathbf{r}(t)$ is defined. Let us consider an example:

Example: What is the domain of

$$\mathbf{r}(t) = t\mathbf{i} + \sqrt{t}\mathbf{j} + \ln(t-3)\mathbf{k}\,. \tag{6}$$

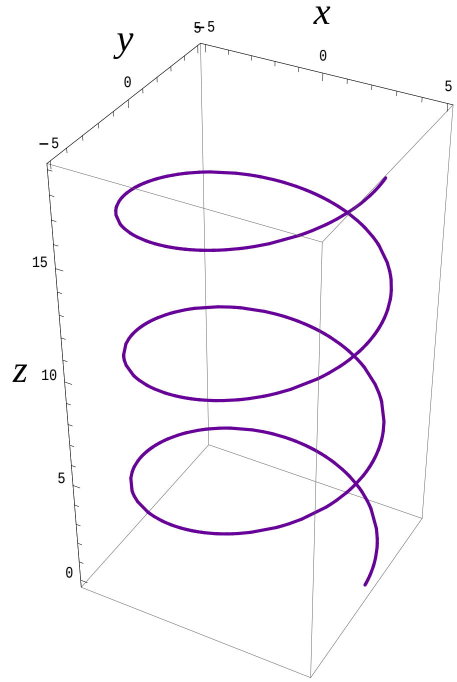

Figure 3: $\mathbf{r}(t) = 5\cos t\mathbf{i} + 5\sin t\mathbf{j} + t\mathbf{k}$

Solution: Where is $\mathbf{r}(t)$ defined? Let's look at it component by component. Firstly, $f(t) = t$ is defined for any real number, and so has the range $(-\infty, \infty)$. Next, $g(t) = \sqrt{t}$ is defined for any non-negative number, so the range is $[0, \infty)$. Finally, the component $h(t) = \ln(t-3)$ is well-defined when the argument of ln is positive; in other words, the range is $(3, \infty)$. So what is the range of $\mathbf{r}(t)$? It is only defined where *all* its components are defined. Therefore its range is $(3, \infty)$.

In this course we will only consider vector-valued functions in two or three dimensions, with values in $t \in \mathbf{R}^2$ and $t \in \mathbf{R}^3$ respectively . Let us look at simple examples. In two dimensions, the parametric equations of a circle of radius 5 are

$$f(t) = 5\cos t\,, \quad g(t) = 5\sin t\,, \tag{7}$$

which shown in Figure 2. we now consider a three-dimensional example with equations

$$f(t) = 5\cos t\,, \quad g(t) = 5\sin t\,, \quad h(t) = t\,, \tag{8}$$

which is instead a spiral (circular helix) and is drawn in Figure 3.

2.1.1 Vector Form of a Line Segment

How do we write a line segment in the form of a parametric equation? Consider Figure 4. If we consider the line that connects the positions $\mathbf{r}_0$ and $\mathbf{r}_1$,

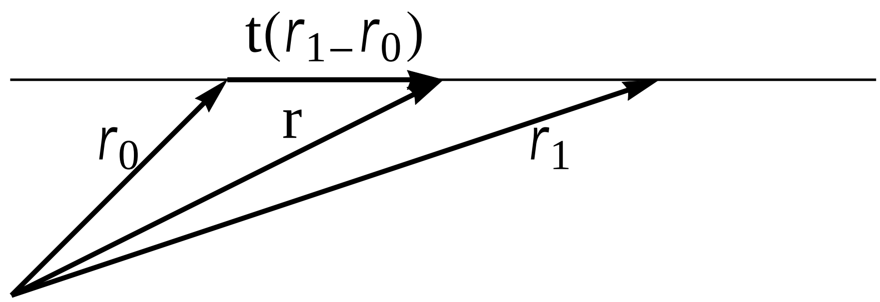

Figure 4: Line segment

we see that to move along this line, we could start at position $\mathbf{r}_0$ and move to other point $\mathbf{r}$ via

$$\mathbf{r} = \mathbf{r}_0 + \mathbf{v}t\,, \tag{9}$$

where $\mathbf{v}$ is the rate at which the position changes (the "velocity") and the parameter t describes how much "time" has elapsed in which the change has occurred. Let $\mathbf{r}_0$ be at $t = 0$ and $\mathbf{r}_1$ be at $t = 1$. Then

$$\mathbf{r}_1 = \mathbf{r}_0 + \mathbf{v}\,, \tag{10}$$

and so

$$\mathbf{v} = \mathbf{r}_1 - \mathbf{r}_0\,. \tag{11}$$

This means we arrive at

$$\begin{aligned} \mathbf{r} &= \mathbf{r}_0 + (\mathbf{r}_1 - \mathbf{r}_0)t \\ &= (1-t)\mathbf{r}_0 + t\mathbf{r}_1\,. \end{aligned} \tag{12}$$

This is the two-point form of a line. Notice that if $0 \le t \le 1$ then it describes the line segment from $\mathbf{r}_0$ to $\mathbf{r}_1$.

2.1.2 Limits and continuity

We define the **limit** of a vector-valued function to be

$$\lim_{t\to a} \mathbf{r}(t) = \mathbf{L}\,, \tag{13}$$

if for $\mathbf{r}(t) = (x(t), y(t), z(t))$

$$\lim_{t\to a} x(t) = L_x\,, \quad \lim_{t\to a} y(t) = L_y\,, \quad \lim_{t\to a} x(t) = L_z\,, \tag{14}$$

with $\mathbf{L} = (L_x, L_y, L_z)$. Using this definition of a limit, we say that a vector-valued function is **continuous** at the point $t = a$ if

$$\lim_{t\to a} \mathbf{r}(t) = \mathbf{r}(a)\,. \tag{15}$$

A vector-valued function is continuous in an interval if it is continuous at

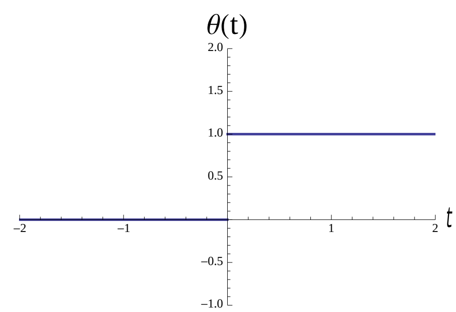

Figure 5: Heaviside step function

all points in the interval. We can see what this means with an example. Clearly, the vector-valued function $\mathbf{r}(t) = (t, t^2, t + t^3)$ is continuous because the limit at $t \to a$ is $\mathbf{r}(a) = (a, a^2, a + a^3)$, and is continuous for all real values of a, i.e. on the interval $(-\infty, \infty)$. However, let us instead consider $\mathbf{r}(t) = (\theta(t), t^2, t + t^3)$, with the new function given by

$$\theta(t) = \begin{cases} 1 & \text{if } t \geq 0 \\ 0 & \text{if } t < 0\,. \end{cases} \tag{16}$$

which is the Heaviside step function. It is shown in Figure 5. Notice that it is continuous on the intervals $[0, \infty)$ and $(-\infty, 0)$, but if we consider the interval $[-2, 0]$, it is discontinuous at the point $t = 0$. To see this, we notice that the limit at $t = 0$ in this interval must be 0 since we must take the limit from below. However, the value at $t = 0$ is 1. Therefore, because one of the components of $\mathbf{r}(t) = (\theta(t), t^2, t + t^3)$ is discontinuous $\mathbf{r}(t)$ is discontinuous.

2.1.3 Derivatives

We now define **derivatives** of vector-valued functions using limits. In order to be differentiable, the vector-valued function must be continuous, but the converse does not hold. The derivative is defined as

$$\mathbf{r}'(t) = \lim_{t \to h} \frac{\mathbf{r}(t+h) - \mathbf{r}(t)}{h}\,, \tag{17}$$

provided the limit exists. Clearly, it exists only when the function is continuous. This is shown in Figure 6. Notice that the derivative $\mathbf{r}'(t)$ is **tangent** to the curve traced out by $\mathbf{r}(t)$, and points in the direction of increasing t. In mechanics, $\mathbf{r}'(t) = \mathbf{v}(t)$. Alternative notations include $\mathbf{r}'$, $\frac{d}{dt}\mathbf{r}(t)$, and $\frac{d\mathbf{r}}{dt}$. Let's look at an example:

$$\begin{aligned} \mathbf{r}(t) &= e^{t^3}\mathbf{i} + \sqrt{1+t^2}\mathbf{j} - \sin t\,\mathbf{k}\,, \\ \Rightarrow \quad \mathbf{r}'(t) &= 3t^2 e^{t^3}\mathbf{i} + \frac{t}{\sqrt{1+t^2}}\mathbf{j} - \cos t\,\mathbf{k}\,. \end{aligned} \tag{18}$$

Let us recall some properties of derivatives that apply to vector-valued functions :

Rules for Differentiation

1. $\frac{d}{dt}\mathbf{c} = 0\,,$
2. $\frac{d}{dt}(k\mathbf{r}) = k\frac{d}{dt}\mathbf{r}\,,$
3. $\frac{d}{dt}(\mathbf{r}_1 + \mathbf{r}_2) = \mathbf{r}_1' + \mathbf{r}_2'\,,$
4. $\frac{d}{dt}(\mathbf{r}_1 - \mathbf{r}_2) = \mathbf{r}_1' - \mathbf{r}_2'\,,$
5. $\frac{d}{dt}(a\mathbf{r}_1 + b\mathbf{r}_2) = a\mathbf{r}_1' + b\mathbf{r}_2'\,,$
6. $\frac{d}{dt}(f(t)\mathbf{r}) = f'(t)\mathbf{r} + f(t)\mathbf{r}'\,,$

where $\mathbf{c}$ is a constant vector and a, b, k are constants, and f is any function of t. The last of these is the chain rule.

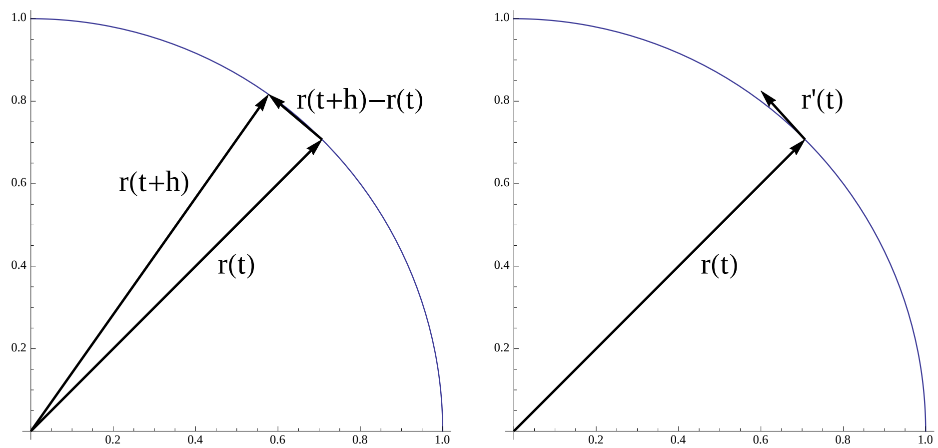

Figure 6: Derivatives of vectors

2.1.4 Tangent Lines and Tangent Vectors

We mentioned in the last section that the derivative of a vector-valued function is tangent to the vector-valued function at that point. From this, we define the **tangent line** to $\mathbf{r}(t)$ at t_0 to be the line parallel to the derivative $\mathbf{r}'(t_0)$. The equation for the tangent line is

$$\mathbf{R}(t) = \mathbf{r}_0 + \mathbf{v}_0 t \,, \tag{19}$$

with $\mathbf{r}_0 = \mathbf{r}(t_0)$ and $\mathbf{v}_0 = \mathbf{r}'(t_0)$. This is clearly the vector form of a line segment.

Example: Find the tangent line of a circular helix with the equation

$$\mathbf{r}(t) = \rho \cos t \, \mathbf{i} + \rho \sin t \, \mathbf{j} + ct \, \mathbf{k} \,, \tag{20}$$

Solution: The derivative is

$$\mathbf{r}'(t) = -\rho \sin t \, \mathbf{i} + \rho \cos t \, \mathbf{j} + c \, \mathbf{k} \,. \tag{21}$$

Let us calculate the tangent line at $t = \pi$. We then have

$$\mathbf{r}_0 = \mathbf{r}(\pi) = -\rho \, \mathbf{i} + \pi c \, \mathbf{k} \,, \tag{22}$$

and

$$\mathbf{v}_0 = \mathbf{r}'(\pi) = -\rho \, \mathbf{j} + c \, \mathbf{k} \,. \tag{23}$$

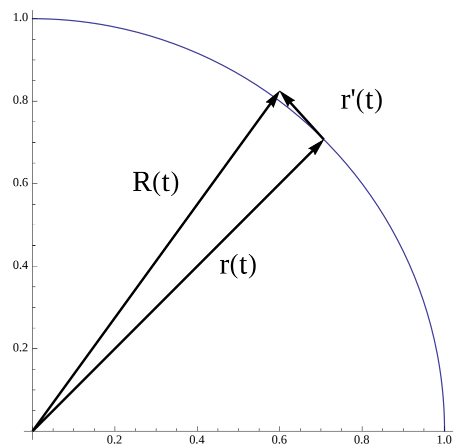

Figure 7: Tangent Line

Therefore the tangent line has the equation

$$\begin{aligned}\mathbf{R}(t) &= (-\rho\,\mathbf{i} + \pi c\,\mathbf{k}) + (-\rho\,\mathbf{j} + c\,\mathbf{k})t \\ &= -\rho\,\mathbf{i} - \rho t\,\mathbf{j} + c(t+\pi)\,\mathbf{k}\,.\end{aligned} \tag{24}$$

2.1.5 Derivatives involving vectors: dot and cross products

Recall that for vectors $\mathbf{r}_1 = (x_1, y_1, z_1)$ and $\mathbf{r}_2 = (x_2, y_2, z_2)$ the dot product is given by

$$\mathbf{r}_1 \cdot \mathbf{r}_2 = x_1 x_2 + y_1 y_2 + z_1 z_2\,, \tag{25}$$

and the magnitude is given by $||\mathbf{r}|| = \sqrt{\mathbf{r}\cdot\mathbf{r}}$. The cross product is given by

$$\mathbf{r}_1 \times \mathbf{r}_2 = (y_1 z_2 - z_1 y_2)\mathbf{i} + (z_1 x_2 - x_1 z_2)\mathbf{j} + (x_1 y_2 - y_1 x_2)\mathbf{k}\,, \tag{26}$$

and its magnitude is

$$||\mathbf{r}_1 \times \mathbf{r}_2|| = ||\mathbf{r}_1|| \cdot ||\mathbf{r}_2|| \sin\theta\,, \tag{27}$$

where θ is the angle between the vectors. Some useful identities are

$$\mathbf{r}_1 \cdot \mathbf{r}_2 = \mathbf{r}_2 \cdot \mathbf{r}_1\,, \quad \mathbf{r}_1 \times \mathbf{r}_2 = -\mathbf{r}_2 \times \mathbf{r}_1\,. \tag{28}$$

Then, applying the derivatives to these products gives

$$\begin{aligned} \frac{d}{dt}(\mathbf{r}_1 \cdot \mathbf{r}_2) &= \mathbf{r}_1' \cdot \mathbf{r}_2 + \mathbf{r}_1 \cdot \mathbf{r}_2' \,, \\ \frac{d}{dt}(\mathbf{r}_1 \times \mathbf{r}_2) &= \mathbf{r}_1' \times \mathbf{r}_2 + \mathbf{r}_1 \times \mathbf{r}_2' \,. \end{aligned} \tag{29}$$

Finally in this section, we use these rules to prove a theorem. You probably remember that the tangents you met previously are perpendicular (normal) to the curve. Is that true here? The theorem below says yes in certain conditions.

Theorem 1.1 If $\mathbf{r}(t)$ is a real vector-valued function with constant magnitude $||\mathbf{r}(t)||$, then $\mathbf{r} \cdot \mathbf{r}' = 0$, which means that $\mathbf{r}'$ is perpendicular to $\mathbf{r}$, i.e. $\mathbf{r} \perp \mathbf{r}'$.

Proof

$$\begin{aligned} & ||\mathbf{r}||^2 = \mathbf{r} \cdot \mathbf{r} \\ \Rightarrow \quad & 0 = \frac{d}{dt}(\mathbf{r} \cdot \mathbf{r}) = \mathbf{r}' \cdot \mathbf{r} + \mathbf{r} \cdot \mathbf{r}' = 2\mathbf{r}' \cdot \mathbf{r} \,, \end{aligned} \tag{30}$$

as required.

2.1.6 Definite Integrals of Vector-Valued Functions

Let $\mathbf{r}(t)$ be a continuous (not necessarily differentiable) function on an interval $a \leq t \leq b$. The **definite integral** is defined as

$$\begin{aligned} \int_a^b \mathbf{r}(t)\,dt &= \lim_{\max \Delta t_k \to 0} \sum_{k=1}^{N} \mathbf{r}(t_k)\Delta t_k \\ &= \int_a^b x(t)\,dt\,\mathbf{i} + \int_a^b y(t)\,dt\,\mathbf{j} + \int_a^b z(t)\,dt\,\mathbf{k} \,. \end{aligned} \tag{31}$$

If desired, this can be understood by considering Figure 8. However, the details are not necessary, and the formula (31) suffices for our purposes. This gives rise to the following useful properties:

Rules of Integration

1. $\int_a^b (k\mathbf{r}(t))\,dt = k \int_a^b \mathbf{r}(t)\,dt \,,$

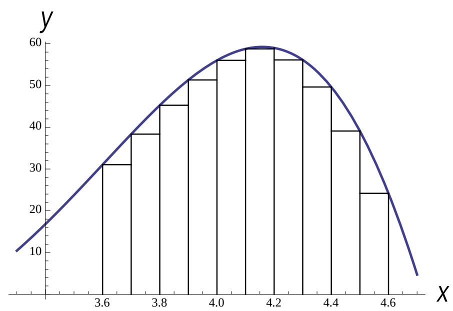

Figure 8: Riemann Integral

2. $\int_a^b (c\,\mathbf{r}_1(t) + d\,\mathbf{r}_2(t))\,dt = c\int_a^b \mathbf{r}_1(t)\,dt + d\int_a^b \mathbf{r}_2(t)\,dt\,.$

Here k, c, and d are constants. Also recall that if the graph f lies above the graph g, then the area between the graphs is given by

$$\text{Area} = \int_a^b \big(f(x) - g(x)\big)\,dx\,. \tag{32}$$

Also recall that the volume of a solid that has cross-sectional area $A(x)$ has volume

$$\text{Volume} = \int_a^b A(x)\,dx\,. \tag{33}$$

Let us consider an example.
Example: Find the integral of $\mathbf{r}(t) = t^3\mathbf{i} + \sqrt{t}\mathbf{j} - \sin\frac{\pi t}{2}\,\mathbf{k}$ with limits 0 and 2.
Solution:

$$\begin{aligned}
&\int_0^2 \left(t^3\mathbf{i} + \sqrt{t}\mathbf{j} - \sin\frac{\pi t}{2}\,\mathbf{k}\right)dt \\
&= \left[\frac{1}{4}t^4\mathbf{i} + \frac{2}{3}t^{3/2}\mathbf{j} + \frac{2}{\pi}\cos\frac{\pi t}{2}\,\mathbf{k}\right]\Bigg|_0^2 \\
&= \left[4\mathbf{i} + \frac{4\sqrt{2}}{3}\mathbf{j} - \frac{2}{\pi}\,\mathbf{k}\right] - \left[0\mathbf{i} + 0\mathbf{j} + \frac{2}{\pi}\,\mathbf{k}\right] \\
&= 4\mathbf{i} + \frac{4\sqrt{2}}{3}\mathbf{j} - \frac{4}{\pi}\,\mathbf{k}\,.
\end{aligned} \tag{34}$$

We made use of the fact that $\cos 0 = 1$ and $\cos\pi = -1$.

Next we define the **antiderivative** for a vector-valued function $\mathbf{r}(t)$, which is itself a vector-valued function $\mathbf{R}(t)$ which is given by

$$\mathbf{R}'(t) = \mathbf{r}(t)\,. \tag{35}$$

Hence,

$$\int \mathbf{r}(t)\,dt = \mathbf{R}(t) + \mathbf{C}\,. \tag{36}$$

Again, let's look at an example,

$$\begin{aligned} &\int \left(\frac{1}{t-1}\,\mathbf{i} + \cos 2t\,\mathbf{j}\right) dt \\ &= \log|t-1|\,\mathbf{i} + \frac{1}{2}\sin 2t\,\mathbf{j} + \mathbf{C}\,. \end{aligned} \tag{37}$$

There are the following properties

1. $\frac{d}{dt}\int \mathbf{r}(t)\,dt = \mathbf{r}(t)\,,$
2. $\int \mathbf{r}'(t)\,dt = \mathbf{r}(t) + \mathbf{C}\,,$

The last thing we consider in this section is **The Fundamental Theorem of Calculus**, which states

$$\int_a^b \mathbf{r}(t)\,dt = \mathbf{R}(t)\Big|_a^b = \mathbf{R}(b) - \mathbf{R}(a)\,. \tag{38}$$

Let us return to the previous example, which gives

$$\begin{aligned} &\int_2^3 \left(\frac{1}{t-1}\,\mathbf{i} + \cos 2t\,\mathbf{j}\right) dt \\ &= \left(\log|t-1|\,\mathbf{i} + \frac{1}{2}\sin 2t\,\mathbf{j}\right)\Bigg|_2^3 \\ &= (\log 2 - \log 1)\,\mathbf{i} + \frac{1}{2}(\sin 6 - \sin 4)\mathbf{j} \\ &= \log 2\,\mathbf{i} + \frac{1}{2}(\sin 6 - \sin 4)\mathbf{j}\,, \end{aligned} \tag{39}$$

where of course $\log 1 = 0$.

2.1.7 Arc Length and Changing Parameters

We say that a curve is **smoothly parameterised** by $\mathbf{r}(t)$, or that $\mathbf{r}(t)$ is a **smooth function** of t if $\mathbf{r}'(t)$ exists, is continuous and that $\mathbf{r}'(t) \neq 0$ for all t. If a function is smooth, we can calculate its **arc length**, which is given by

$$L = \int_a^b \left|\left|\frac{d\mathbf{r}}{dt}\right|\right| dt = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}\, dt\,. \tag{40}$$

As an example, lets find the arc length of $\mathbf{r}(t) = 5\cos t\mathbf{i} + 5\sin t\mathbf{j} + t\mathbf{k}$ from 0 to $\frac{\pi}{2}$. Then $\mathbf{r}'(t) = -5\sin t\mathbf{i} + 5\cos t\mathbf{j} + \mathbf{k}$, and

$$||\mathbf{r}'(t)|| = \sqrt{(-5\sin t)^2 + (5\cos t)^2 + 1} = \sqrt{26}\,, \tag{41}$$

which gives arc length

$$L = \int_0^{\pi/2} \sqrt{26}\, dt = \sqrt{26}\frac{\pi}{2}\,. \tag{42}$$

2.1.7.1 Parameterising a curve using arc length

In order to use the arc length to parameterise a curve, we select a point to be our reference point, P, and choose an orientation so that one direction is positive and the other negative. Then, the "signed" arc length parameterises the curve via

$$x = x(s)\,, \quad y = y(s)\,, \quad z = z(s)\,. \tag{43}$$

In other words, we view the position on the curve as a function of its arc length ("distance") from the reference point, P. This is known as **arc length parameterisation**. To understand how it works, let us consider an example of a circle of radius 5. It has the parametric equation

$$\mathbf{r}(t) = 5\cos t\,\mathbf{i} + 5\sin t\,\mathbf{j}\,, \quad 0 \le t \le 2\pi\,, \tag{44}$$

in which t plays the role of a radian angle measured from the point O on the x-axis to an arbitrary point $P(x, y)$, as shown in Figure 9. The arc length is

$$s \equiv L = \int_0^t \sqrt{(-5\sin t)^2 + (5\cos t)^2}\, dt = 5t\,, \tag{45}$$

and therefore

$$s = 5t\,, \quad \text{or} \quad t = s/5\,. \tag{46}$$

As a result, the circle is now parameterised by

$$x = 5\cos(s/5)\,, \quad y = 5\sin(s/5)\,, \quad 0 \le s \le 10\pi\,. \tag{47}$$

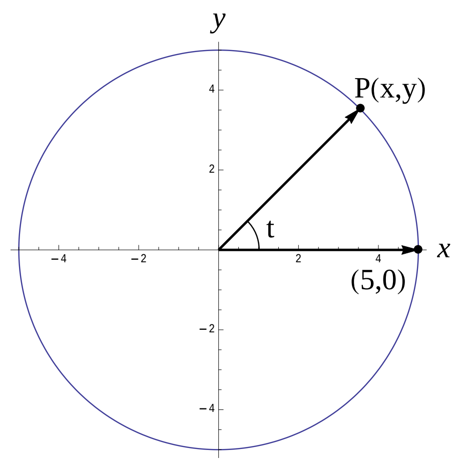

Figure 9: Arc length parameterisation of a circle

2.1.7.2 Change of parameter

More generally, we can change the parameter of a curve using the chain rule. If $\mathbf{r}(t)$ is a vector-valued function that is differentiable with respect to t, we can change parameter to τ using $t = g(\tau)$, where g is differentiable with respect to τ. Then $\mathbf{r}(g(\tau)$ is differentiable with respect to τ and we have

$$\frac{d\mathbf{r}}{d\tau} = \frac{d\mathbf{r}}{dt}\frac{dt}{d\tau}\,. \tag{48}$$

If $g(\tau)$ is smooth, then the change is called a **smooth change of parameter**. Also, if $\frac{dt}{d\tau} > 0$ for all τ, then it is a **positive change of parameter**. Alternatively, if $\frac{dt}{d\tau} < 0$ for all τ, then it is a **negative change of parameter**. Using these notations, we state that if C is the graph of a smooth vector-valued function $\mathbf{r}(t)$, with some reference point $\mathbf{r}(t_0)$, then the arc length parameter is given by a positive change of parameter and is found from the formula

$$s = \int_{t_0}^{t} \left|\left|\frac{d\mathbf{r}}{du}\right|\right| du = \int_{t_0}^{t} \sqrt{\left(\frac{dx}{du}\right)^2 + \left(\frac{dy}{du}\right)^2 + \left(\frac{dz}{du}\right)^2}\, du\,. \tag{49}$$

For an example, let us consider the vector-valued function $\mathbf{r}(u) = 2\cos u\,\mathbf{i} + 2\sin u\,\mathbf{j} + u\,\mathbf{k}$, where u is some parameter replacing t in order to avoid poor notation in the integral. We will use $t_0 = 0$ as our reference point. Then

$$\left|\left|\frac{d\mathbf{r}}{du}\right|\right| = \sqrt{5}\,, \tag{50}$$

and so

$$s = \int_0^t \sqrt{5}du = \sqrt{5}t\,. \tag{51}$$

Therefore, the reparameterisation give us

$$\mathbf{r}(s) = 2\cos\left(\frac{s}{\sqrt{5}}\right)\mathbf{i} + 2\sin\left(\frac{s}{\sqrt{5}}\right)\mathbf{j} + \frac{s}{\sqrt{5}}\mathbf{k}\,. \tag{52}$$

Important note: Although, this "new" variable u might seem confusing, it really is the same thing as t. It is a dummy variable representing t for the purpose of integration. Otherwise we would end up with equally confusing and mathematical incorrect expressions such as $\int_0^t ||d\mathbf{r}/dt||dt$, which obviously we want to avoid.

Although we can define many different reparameterisations, the arc length parameterisation has some special properties. These are

(a) If $\mathbf{r}(t)$ is a vector-valued function with parameter t, with arc length parameter s,

$$\left|\left|\frac{d\mathbf{r}}{dt}\right|\right| = \frac{ds}{dt}\,. \tag{53}$$

(b) If $\mathbf{r}(t)$ is a vector-valued function and arc length parameter s, the tangent vector for any s has length given by

$$\left|\left|\frac{d\mathbf{r}}{ds}\right|\right| = 1\,. \tag{54}$$

(a) If $\mathbf{r}(t)$ is a vector-valued function with $||d\mathbf{r}/dt|| = 1$ for any value of parameter t, then $s = t - t_0$ is an arc length parameter with reference point t_0.

The proof of these is simple. For (a), apply the Fundamental Theorem of Calculus to (49). Then the antiderivative is s and therefore, ds/dt is the argument of the integral, which is $||d\mathbf{r}/dt||$, and we get the result. (b) comes directly from (a) if we take $t = s$. Finally, (c) comes from setting $||d\mathbf{r}/dt|| = 1$ in (49), which gives

$$s = \int_{t_0}^{t} \left|\left|\frac{d\mathbf{r}}{du}\right|\right| du = \int_{t_0}^{t} du = u|_{t_0}^{t} = t - t_0 \,, \tag{55}$$

as required.

2.1.8 Tangent, Normal and Binormal Vectors

We previously defined the tangent vector for the curve of a vector-valued function to be the vector $\mathbf{r}'$. We often instead want the **unit tangent vector** which is the unit vector along the tangent line, i.e.

$$\mathbf{T} = \frac{\mathbf{v}(t)}{||\mathbf{v}(t)||} = \frac{\mathbf{r}'(t)}{||\mathbf{r}'(t)||} \,, \tag{56}$$

where of course $||\mathbf{v}(t)||$ is the magnitude of the vector. This is important to allow us to define the normal vector.

2.1.8.1 Normal vectors

Recall that we said that if a vector-valued function has a constant norm $||\mathbf{r}||$, then the tangent (and hence unit tangent) vector to the curve C is orthogonal to $\mathbf{r}$. If we apply the same reasoning to $\mathbf{T}$, which has constant norm 1, we see that $\mathbf{T}'$ must be orthogonal to $\mathbf{T}$. We say that $\mathbf{T}'$ is **normal** to C, and, provided $\mathbf{T}' \neq 0$, we define the **(principal) unit normal vector** to C at t as

$$\mathbf{N}(t) = \frac{\mathbf{T}'(t)}{||\mathbf{T}'(t)||} \,. \tag{57}$$

Note that the condition $\mathbf{T}' \neq 0$ means that we cannot define the normal vector for straight lines. An indication of these vectors is given in Figure 10.

Example: Find the unit tangent and unit normal vectors to the curve

$$\mathbf{r}(t) = 3\cos t\mathbf{i} + 3\sin t\mathbf{j} + 4t\mathbf{k} \,.$$

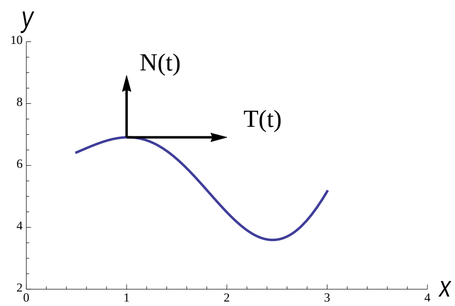

Figure 10: Unit tangent and unit normal vectors.

Solution: We first find the tangent vector,

$$\mathbf{r}'(t) = -3\sin t\mathbf{i} + 3\cos t\mathbf{j} + 4\mathbf{k}\,,$$

which has magnitude

$$||\mathbf{r}'(t)|| = \sqrt{(-3\sin t)^2 + (3\cos t)^2 + 4^2} = 5\,,$$

which gives us the unit tangent vector

$$\mathbf{T}(t) = -\frac{3}{5}\sin t\mathbf{i} + \frac{3}{5}\cos t\mathbf{j} + \frac{4}{5}\mathbf{k}\,.$$

Then

$$\mathbf{T}'(t) = -\frac{3}{5}\cos t\mathbf{i} - \frac{3}{5}\sin t\mathbf{j}\,,$$

which has magnitude

$$||\mathbf{T}'(t)|| = \sqrt{\left(-\frac{3}{5}\cos t\right)^2 + \left(-\frac{3}{5}\sin t\right)^2} = \frac{3}{5}\,,$$

which gives us the unit normal vector

$$\mathbf{N}(t) = -\cos t\mathbf{i} - \sin t\mathbf{j}\,.$$

Note that this vector is the parametric form of a circle in a plane with a minus sign. Hence the normal vector is parallel to the xy-plane and points from the curve (spiral) towards the z-axis, i.e. in the opposite direction to the vector that would describe a circle.

When the curve is parameterised by its arc length parameter, the calculation of these vectors is simpler. We have the formula

$$\begin{aligned}\mathbf{T}(s) &= \mathbf{r}'(s)\,, \\ \mathbf{N}(s) &= \frac{\mathbf{r}''(s)}{||\mathbf{r}''(s)||}\,.\end{aligned} \tag{58}$$

Example: Calculate the unit tangent and unit normal vectors of

$$\mathbf{r}(t) = 3\cos t\mathbf{i} + 3\sin t\mathbf{j}\,, \quad (0 \leq t \leq 2\pi)\,.$$

Solution: The arc length parameter is

$$s = \int_0^t 3du = 3t\,,$$

and therefore

$$\mathbf{r}(s) = 3\cos s/3\mathbf{i} + 3\sin s/3\mathbf{j}\,.$$

We then find the tangent vector,

$$\mathbf{T}(s) = \mathbf{r}'(s) = -\sin s/3\mathbf{i} + \cos s/3\mathbf{j}\,.$$

Then

$$\mathbf{r}''(s) = -\frac{1}{3}\cos s/3\mathbf{i} - \frac{1}{3}\sin s/3\mathbf{j}\,,$$

which has magnitude

$$||\mathbf{r}''(s)|| = \sqrt{\left(-\frac{1}{3}\cos s/3\right)^2 + \left(-\frac{1}{3}\sin s/3\right)^2} = \frac{1}{3}\,,$$

which gives us the unit normal vector

$$\mathbf{N}(s) = -\cos s/3\mathbf{i} - \sin s/3\mathbf{j}\,.$$

2.1.8.2 Binormal vectors

The **binormal vector** is defined by

$$\mathbf{B}(t) = \mathbf{T}(t) \times \mathbf{N}(t)\,. \tag{59}$$

By definition it is orthogonal to both **T** and **N**, and since both these vectors are unit vectors, **B** is also a unit vector, since

$$||\mathbf{T}(t) \times \mathbf{N}(t)|| = ||\mathbf{T}||||\mathbf{N}|| \sin \pi/2 = 1\,. \tag{60}$$

These three vectors define three mutually perpendicular planes at any point

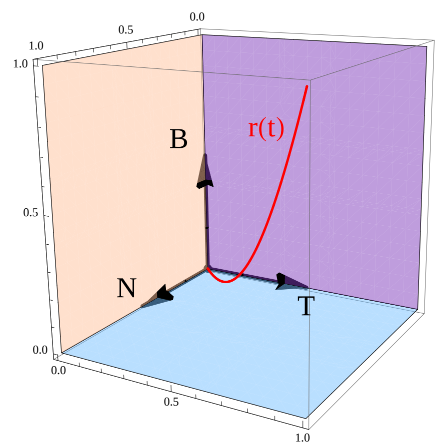

Figure 11: Normal, binormal and tangent vectors, and the planes containing them.

on a curve C of a vector-valued function $\mathbf{r}(t)$. These are the **TB-plane** or **rectifying plane**, the **TN-plane** or **osculating plane**, and the **NB-plane** or **normal plane**. This is shown in Figure 11. This defines a coordinate system known as the **TBN-frame** or **Frenet frame**, which is a frame that has its origin move along the curve C, with its axes rotating as we move along the curve (since **T**, **N** and **B** rotate). We can express **B** directly in terms of the original vector-valued function by

$$\mathbf{B}(t) = \frac{\mathbf{r}'(t) \times \mathbf{r}''(t)}{||\mathbf{r}'(t) \times \mathbf{r}''(t)||}\,, \tag{61}$$

which in the situation where we use the arc length parameter, becomes

$$\mathbf{B}(s) = \frac{\mathbf{r}'(s) \times \mathbf{r}''(s)}{||\mathbf{r}''(s)||}\,. \tag{62}$$

This is a good moment to pause to think again about why we might want to use arc length parameterisation. We have already seen that it simplifies the simplifies the expressions for the unit tangent, unit normal and binormal vectors. This is itself useful, but it is in the definition of curvature that it becomes vital.

2.1.9 Curvature

For a curve C of a vector-valued function $\mathbf{r}$, the unit tangent vector $\mathbf{T}$ is a measure of how quickly we move along C. If we then consider the derivative $d\mathbf{T}/ds$, this is a measure of how quickly the motion changes direction or "curves" at that point. Therefore, we define the **curvature** of C to be

$$\kappa(s) = \left|\left|\frac{d\mathbf{T}}{ds}\right|\right| = ||\mathbf{r}''(s)||\,, \tag{63}$$

which is a number which tells us how much a curve bends. Of course, this is a bit simplistic: a curve in three dimensions can bend in three different directions, so how can one number give us all this information? The answer is that it doesn't, and in fact, we should also consider $d\mathbf{N}/ds$ and $d\mathbf{B}/ds$ to gives us a full picture. The concept of $\mathbf{T}$ changing along the curve is represented in Figure 12.

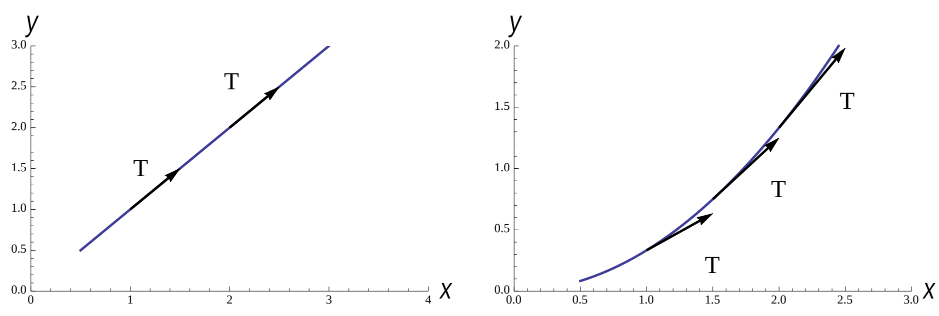

Figure 12: $\mathbf{T}$ along a line and a generic curve.

Example: Find the curvature of

(a) The circle $\mathbf{r}(s) = 3\cos s/3\mathbf{i} + 3\sin s/3\mathbf{j}\,, \quad (0 \le s \le 6\pi)\,.$

(b) The line $\mathbf{r}(s) = \mathbf{r}_0 + s\mathbf{v}\,.$

Solution: We apply the formula for the curvature:

(a) The second derivative is

$$\mathbf{r}''(s) = -\frac{1}{3}\cos s/3\mathbf{i} - \frac{1}{3}\sin s/3\mathbf{j}\,,$$

and so the curvature is

$$\kappa(s) = \sqrt{\left(-\frac{1}{3}\cos s/3\right)^2 + \left(-\frac{1}{3}\sin s/3\right)^2} = \frac{1}{3}\,.$$

(b) The first derivative is

$$\mathbf{r}'(s) = 0 + \mathbf{v} = \mathbf{v}\,,$$

since $\mathbf{r}_0$ and $\mathbf{v}$ are constants, and the second derivative will vanish as a result:

$$\mathbf{r}''(s) = 0\,.$$

This implies the curvature is zero:

$$\kappa(s) = 0\,.$$

More generally, we find that for a circle of radius a, the curvature is $1/a$ and so the the larger a circle, the smaller the curvature, which makes sense intuitively. Also, any line has zero curvature, which again makes sense.
If we want to use more a more general parameter t, we can write the curvature as

$$\begin{aligned} \kappa(t) &= \frac{||\mathbf{T}'(t)||}{||\mathbf{r}'(t)||}\,, \\ \text{or } \kappa(t) &= \frac{||\mathbf{r}'(t) \times \mathbf{r}''(t)||}{||\mathbf{r}'(t)||^3}\,, \end{aligned} \tag{64}$$

but the expression is much simpler for the arc length parameter. We will not prove these, but proofs can be found in section 12.5 of the textbook. Often, the second of these formulae is more practical to use. These are necessary when calculating the arc length parameter is complicated. For example if the argument in the square root in equation (49) cannot be reduced to a square it might be very difficult to integrate the result. In these cases the formulae for the general parameter are more appropriate.

Example: Consider the vector-valued function $\mathbf{r} = 2\cos t\mathbf{i} + 3\sin t\mathbf{j}$ for $0 \leq t \leq 2\pi$. Find the curvature at the points $t = \pi/2$ and $t = \pi$.

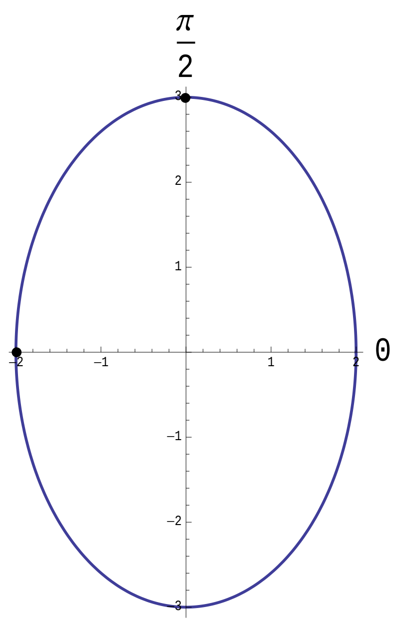

Figure 13: The graph of $\mathbf{r} = 2\cos t\mathbf{i} + 3\sin t\mathbf{j}$.

Solution: The graph of this vector-valued function is an ellipse, shown in Figure 13. You should recognise it as the polar coordinates of an ellipse. Therefore we are looking for the curvature at the endpoints of the major and minor axes. Since the curvature equations are three-dimensional, we view the ellipse to sit in the xy-plane of a three-dimensional coordinate system and write

$$\mathbf{r} = 2\cos t\mathbf{i} + 3\sin t\mathbf{j} + 0\mathbf{k}\,.$$

Then

$$\mathbf{r}' = -2\sin t\mathbf{i} + 3\cos t\mathbf{j}\,,$$
$$\mathbf{r}'' = -2\cos t\mathbf{i} - 3\sin t\mathbf{j}\,,$$

and we find the cross-product of these

$$\mathbf{r}' \times \mathbf{r}'' = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -2\sin t & 3\cos t & 0 \\ -2\cos t & -3\sin t & 0 \end{vmatrix} = \left[6\sin^2 t + 6\cos^2 t\right]\mathbf{k} = 6\mathbf{k}\,.$$

In other words, this vector, which is parallel to the binomial vector points along the z axis. We now need its norm, which is obviously 6. We also need the norm of $\mathbf{r}'$, which is

$$||\mathbf{r}'|| = \sqrt{(-2\sin t)^2 + (3\cos t)^2} = \sqrt{4\sin^2 t + 9\cos^2 t}\,.$$

This is the reason we would be unwise to use the arc length parameter here since this square root appears in the definition of the arc length parameter. Obviously this makes life difficult, so instead we find the curvature via (64) to find

$$\kappa(t) = \frac{||\mathbf{r}'(t) \times \mathbf{r}''(t)||}{||\mathbf{r}'(t)||^3} = \frac{6}{\left(4\sin^2 t + 9\cos^2 t\right)^{3/2}}\,.$$

We now can find the particular values we are interested in

$$\kappa(\pi/2) = \frac{6}{(4)^{3/2}} = \frac{3}{4}\,,$$
$$\kappa(\pi) = \frac{6}{(9)^{3/2}} = \frac{2}{9}\,.$$

Notice that the curvature is greater at the end of the major axis than at the end of the minor axis. Between this points the curvature is between these values as we get a mixture of the sin and cos terms.

2.1.9.1 Radius of Curvature

Our previous example showed us that the curvature of an ellipse varies as we go round the ellipse. This is not surprising, but it highlights that the circle is special with a constant curvature given by $1/a$, where a is the radius of the circle. We can make use of this fact to define **osculating circle** or **circle of curvature** at a point P on a curve to be the circle with the same curvature as the curve at P that also shares a tangent line, and lies on the concave side of the curve. Hence, the circle has radius $\rho = \frac{1}{\kappa}$. ρ is called the **radius of curvature** and the centre of the osculating circle is called the **centre of curvature**. This is shown in Figure 14.

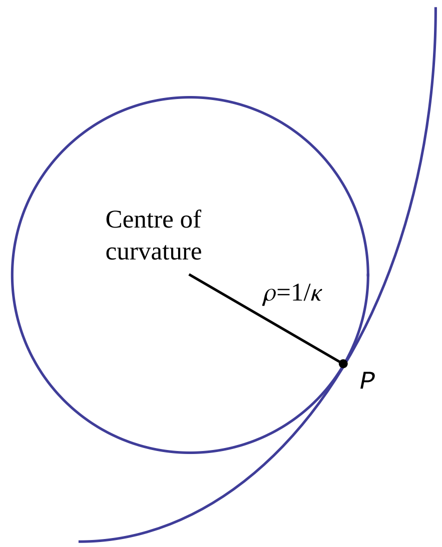

Figure 14: The centre of curvature and radius of curvature for a curve.

2.1.9.2 Curvature in two dimensions

In two dimensions, we can use an angle, ϕ defined with respect to the x-axis to parameterise curvature. You might notice that this is essentially the same idea as a polar angle for polar coordinates. This is shown in Figure 15. We

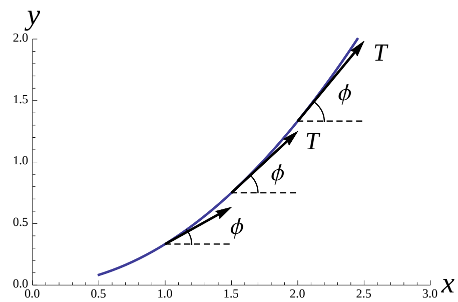

Figure 15: Curvature in two dimensions.

then express the tangent vector along the $\mathbf{i}$ and $\mathbf{j}$ directions as

$$\mathbf{T} = \cos\phi\mathbf{i} + \sin\phi\mathbf{j}\,, \tag{65}$$

which of course gives us

$$\frac{d\mathbf{T}}{d\phi} = -\sin\phi\,\mathbf{i} + \cos\phi\,\mathbf{j}\,. \tag{66}$$

The key point is that we can relate this to the derivative with respect to the arc length parameter using the chain rule. We find

$$\frac{d\mathbf{T}}{ds} = \frac{d\mathbf{T}}{d\phi}\frac{d\phi}{ds}\,. \tag{67}$$

Therefore the curvature becomes

$$\kappa(s) = \left\|\frac{d\mathbf{T}}{ds}\right\| = \left|\frac{d\phi}{ds}\right|\left\|\frac{d\mathbf{T}}{d\phi}\right\| = \sqrt{(-\sin\phi)^2 + (\cos\phi)^2}\left|\frac{d\phi}{ds}\right| = \left|\frac{d\phi}{ds}\right|, \tag{68}$$

which we present compactly as

$$\kappa(s) = \left|\frac{d\phi}{ds}\right|, \tag{69}$$

This is the rate of change of ϕ with respect to s. In other words, in two dimensions, curvature can be interpreted as how quickly the polar coordinate changes.

2.1.10 Mechanics

You might have noticed some striking similarities between what we have been doing and your experience in Newtonian mechanics. This is not a coincidence. Although many mechanics problems are of linear motion or circular motion, and vectors tend to be treated separately when considering work, moments and so on, in fact we should always think three-dimensionally. Think again about arc length. For a straight line is it the same as distance. For circular motion, it is a portion of the orbit. Either way, the instantaneous speed along the curve is ds/dt, which is the rate of change of the arc length. More generally, we can define the **velocity** as

$$\mathbf{v}(t) = \frac{ds}{dt}\mathbf{T}(t)\,, \tag{70}$$

since $\mathbf{T}$ is a unit vector which gives the direction of the rate of change of position. Of course, the acceleration is just the derivative of the velocity,

$\mathbf{a} = \frac{d\mathbf{v}}{dt}$.Therefore, we have the following equations

$$\begin{aligned} \text{velocity } &= \mathbf{v}(t) = \frac{d\mathbf{r}}{dt}\,, \\ \text{acceleration } &= \mathbf{a}(t) = \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{r}}{dt^2}\,, \\ \text{speed } &= ||\mathbf{v}(t)|| = \frac{ds}{dt}\,. \end{aligned} \tag{71}$$

What about displacement? We can simply integrate the velocity to give this, assuming the velocity is known via

$$\Delta\mathbf{r} = \int_{t_1}^{t_2} \mathbf{v}(t)\,dt = \int_{t_1}^{t_2} \frac{d\mathbf{r}(t)}{dt}\,dt = \mathbf{r}(t_2) - \mathbf{r}(t_1)\,. \tag{72}$$

Of course if we already know the position vector (i.e. the vector-valued function for whose graph gives the motion of the particle), we can directly substitute or recognise $\mathbf{r}$ as the antiderivative of $\mathbf{v}$ using the Fundamental Theorem of Calculus. This is something you probably understand very well, but it is helpful to look at it in this mathematical way to help us to understand what all the previous work on vector-valued functions was about. The distance travelled comes from integrating the norm of the velocity (speed) over the time interval, which is

$$s = \int_{t_1}^{t_2} \left|\left|\frac{d\mathbf{r}}{dt}\right|\right| dt = \int_{t_1}^{t_2} ||\mathbf{v}||\,dt\,. \tag{73}$$

Although we use s here, this is not actually an arc length parameter unless we say t_1 is a reference point. The use of s is an unfortunate coincidence due to the fact that in mechanics s is the standard way to represent displacement. This is actually an arc length over an interval.

Returning to the acceleration, you know from circular motion that we should expect both a tangential acceleration and radial acceleration in general. This radial acceleration is in fact along the direction of the normal vector, and using $\mathbf{v} = ds/dt\mathbf{T}$, we have the following decomposition of the acceleration

$$\mathbf{a} = \frac{d^2s}{dt^2}\mathbf{T} + \kappa\left(\frac{ds}{dt}\right)\mathbf{N}\,. \tag{74}$$

We can define two separate components of acceleration

$$\mathbf{a} = a_T\mathbf{T} + a_N\mathbf{N}\,, \tag{75}$$

with

$$a_T = \frac{d^2 s}{dt^2}\,, \quad a_N = \kappa\left(\frac{ds}{dt}\right)\,. \tag{76}$$

We call a_T the **tangential scalar component of acceleration** and a_N the **normal scalar component of acceleration** and $a_T\mathbf{T}$ the **tangential vector component of acceleration** and $a_N\mathbf{N}$ the **normal vector component of acceleration**. Finally, in terms of velocity and acceleration, these are given by

$$a_T = \frac{\mathbf{v}\cdot\mathbf{a}}{||\mathbf{v}||}\,, \quad a_N = \frac{||\mathbf{v}\times\mathbf{a}||}{||\mathbf{v}||}\,, \quad \kappa = \frac{||\mathbf{v}\times\mathbf{a}||}{||\mathbf{v}||^3}\,. \tag{77}$$

Example: Find the velocity, acceleration and tangential and normal accelerations of the motion

$$\mathbf{r}(t) = 2\cos t\mathbf{i} + 3\sin t\cos\theta\mathbf{j} + 3\sin t\sin\theta\mathbf{k}\,,$$

assuming θ is a constant.

Solution: Firstly, this is in fact an ellipse rotated out of the xy-plane by and angle θ about the x-axis. Hence the y and z components have changed. This is shown in Figure 17. It describes the orbit of a planet around a star if the orbit is at an angle to the xy-plane. The velocity is given by

$$\mathbf{v}(t) = -2\sin t\mathbf{i} + 3\cos t\cos\theta\mathbf{j} + 3\cos t\sin\theta\mathbf{k}\,,$$

with magnitude given by

$$||\mathbf{v}|| = \sqrt{(-2\sin t)^2 + (3\cos t\cos\theta)^2 + (3\cos t\sin\theta)^2} = \sqrt{4\sin^2 t + 9\cos^2 t}\,.$$

The acceleration is then

$$\mathbf{a} = -2\cos t\mathbf{i} - 3\sin t\cos\theta\mathbf{j} - 3\sin t\sin\theta\mathbf{k}\,.$$

To find the components of acceleration, we use equation (77). The dot product is

$$\mathbf{v}\cdot\mathbf{a} = \left(2 - 9\sin^2(\theta)\right)\sin(2t)\,,$$

and the cross product is

$$\mathbf{v}\times\mathbf{a} = -6\sin\theta\mathbf{j} + 6\sin\theta\mathbf{k}\,,$$

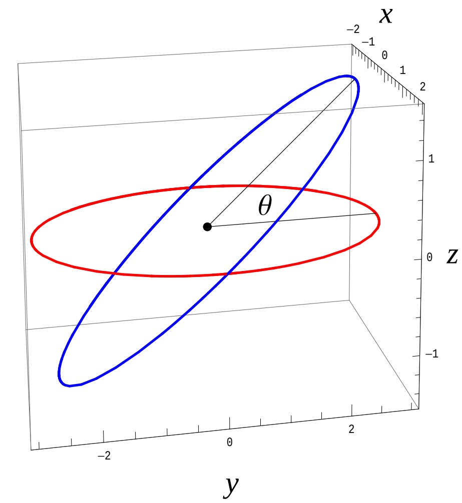

Figure 16: An ellipse in the xy-plane (red) and rotated away from the plane (blue).

with norm

$$||\mathbf{v}\times\mathbf{a}|| = 6\sqrt{2}\sin\theta$$

Note that this is 0 if $\theta = 0$. In other words if the motion is in the xy-plane there is no normal acceleration. The components of acceleration are

$$\begin{aligned} a_T &= \frac{\mathbf{v}\cdot\mathbf{a}}{||\mathbf{v}||} = \frac{\left(2-9\sin^2(\theta)\right)\sin(2t)}{\sqrt{4\sin^2 t + 9\cos^2 t}}\,, \\ a_N &= \frac{||\mathbf{v}\times\mathbf{a}||}{||\mathbf{v}||^3} = \frac{6\sqrt{2}\sin\theta}{\left(4\sin^2 t+9\cos^2 t\right)^{3/2}}\,. \end{aligned} \tag{78}$$

2.1.11 Planetary Orbits - Kepler's Laws

Recall from section 1.6.2 Kepler's Laws of Planetary Motion:

Kepler's Laws of Planetary Motion:

1. **Law of Orbits:** The motion of each planet traces an ellipse with the Sun at one of the foci.

2. **Law of Areas:** The line joining the Sun to the centre of the planet sketches out equal areas in equal times.

3. **Law of Periods:** The square of the orbital period of a planet is proportional to the cube of the semi-major axis of its orbit.

How did Kepler derive these? We now have the tools to understand this. If a particle moves under the influence of a single force acting from a fixed point, we say the particle is moving in a **central force field**, with the force called a **central force**. If we look at the problem in a coordinate system with this fixed point as the origin, the acceleration acts along the radius vector (since the force does) but in the opposite direction. This means

$$\mathbf{r} \times \mathbf{a} = 0\,. \tag{79}$$

As a result, we have another conserved quantity, $\mathbf{l} = \mathbf{r} \times \mathbf{v}$ since

$$\frac{d\mathbf{l}}{dt} = \frac{d}{dt}(\mathbf{r} \times \mathbf{v}) = (\mathbf{v} \times \mathbf{v}) + (\mathbf{r} \times \mathbf{a}) = \mathbf{0} \times \mathbf{0} = \mathbf{0}\,. \tag{80}$$

If you are familiar with the concept of angular momentum, this is essentially the same thing: for angular momentum replace $\mathbf{v}$ with $\mathbf{p}$, the ordinary momentum.

2.1.11.1 Newton's Universal Law of Gravitation

The relevance of central forces is due to Newton's Universal Law of Gravitation, which you should recognise as

$$||\mathbf{F}|| = \frac{GMm}{r^2}\,. \tag{81}$$

If you are not yet familiar with the vector version of this force, note that it acts between two bodies. If $M >> m$, say in the case of the Sun and the Earth, then we treat the centre of M as fixed and the force acts along the radius vector, and points to the centre of M. Hence,

$$\mathbf{F} = ||\mathbf{F}|| \left(-\frac{\mathbf{r}}{||\mathbf{r}||} \right) = ||\mathbf{F}|| \left(-\frac{\mathbf{r}}{r} \right)\,, \tag{82}$$

where $-\frac{\mathbf{r}}{||\mathbf{r}||}$ is a unit vector anti-parallel to $\mathbf{r}$ and

$$\mathbf{F} = -\frac{GMm}{r^3}\mathbf{r}\,. \tag{83}$$

Therefore the acceleration is

$$\mathbf{a} = -\frac{GM}{r^3}\mathbf{r}\,. \tag{84}$$

Take the initial position and velocity to be

$$\mathbf{r}_0 = r_0\mathbf{i}\,, \quad \mathbf{v}_0 = v_0\mathbf{j}\,, \tag{85}$$

which implies

$$\mathbf{l} = r_0\mathbf{i} \times v_0\mathbf{j} = r_0 v_0\mathbf{k}\,, \tag{86}$$

which is constant. If we now define a unit vector

$$\mathbf{u} = \cos\theta\mathbf{i} + \sin\theta\mathbf{j}\,, \tag{87}$$

which follows the planet, we can write the radius vector in polar coordinates using this convenient unit vector:

$$\mathbf{r} = r\cos\theta\mathbf{i} + r\sin\theta\mathbf{j} = r\mathbf{u}\,, \tag{88}$$

and hence

$$\mathbf{a} = -\frac{GM}{r^2}\mathbf{u}\,. \tag{89}$$

Now,

$$\mathbf{v} = \frac{d}{dt}\mathbf{r} = \frac{d}{dt}(r\mathbf{u}) = r\frac{d\mathbf{u}}{dt} + \frac{dr}{dt}\mathbf{u}\,, \tag{90}$$

and since $\mathbf{l}$ is constant,

$$\mathbf{l} = \mathbf{r}\times\mathbf{v} = (r\mathbf{u})\times\left(r\frac{d\mathbf{u}}{dt} + \frac{dr}{dt}\mathbf{u}\right) = r^2\mathbf{u}\times\frac{d\mathbf{u}}{dt} + r\frac{dr}{dt}\mathbf{u}\times\mathbf{u} = r^2\mathbf{u}\times\frac{d\mathbf{u}}{dt}\,, \tag{91}$$

since $\mathbf{u}\times\mathbf{u} = 0$. Since $\mathbf{u}$ depends on θ we must use the chain rule to find its derivative

$$\frac{d\mathbf{u}}{dt} = \frac{d\mathbf{u}}{d\theta}\frac{d\theta}{dt} = (-\sin\theta\mathbf{i} + \cos\theta\mathbf{j})\frac{d\theta}{dt}\,, \tag{92}$$

and so

$$\mathbf{u}\times\frac{d\mathbf{u}}{dt} = \frac{d\theta}{dt}\mathbf{k}\,, \tag{93}$$

and therefore

$$\mathbf{l} = r^2\frac{d\theta}{dt}\mathbf{k}\,. \tag{94}$$

If we now take the cross product between **a** and **l**, we get

$$\begin{aligned}\mathbf{a}\times\mathbf{l} &= -\frac{GM}{r^2}(\cos\theta\mathbf{i}+\sin\theta\mathbf{j})\times\left(r^2\frac{d\theta}{dt}\mathbf{k}\right)\\ &= GM(-\sin\theta\mathbf{i}+\cos\theta\mathbf{j})\frac{d\theta}{dt} = GM\frac{d\mathbf{u}}{dt}\,.\end{aligned} \tag{95}$$

Since **l** is constant, we find

$$\frac{d}{dt}(\mathbf{v}\times\mathbf{l}) = \frac{d\mathbf{v}}{dt}\times\mathbf{l} = \mathbf{a}\times\mathbf{l} = GM\frac{d\mathbf{u}}{dt}\,, \tag{96}$$

which we integrate to find

$$\mathbf{v}\times\mathbf{l} = GM\mathbf{u}+\mathbf{C}\,, \tag{97}$$

where **C** is a constant vector. Since this is a constant, we can find it's value by substituting the values of **v** and **u** at $t=0$, i.e. $\mathbf{v}=v_0\mathbf{j}$ and $\mathbf{u}=\mathbf{i}$ and we get

$$\mathbf{C} = (r_0v_0^2 - GM)\mathbf{i}\,. \tag{98}$$

We will now make use of this to determine the position as a function of θ. Consider $\mathbf{r}\cdot(\mathbf{v}\times\mathbf{l})$ and recall the identity $\mathbf{u}\cdot(\mathbf{v}\times\mathbf{w}) = (\mathbf{u}\times\mathbf{v})\cdot\mathbf{w}$. Then

$$\mathbf{r}\cdot(\mathbf{v}\times\mathbf{l}) = (\mathbf{r}\times\mathbf{v})\cdot\mathbf{l} = \mathbf{l}\cdot\mathbf{l} = r_0^2v_0^2\,. \tag{99}$$

Alternatively,

$$\begin{aligned}\mathbf{r}\cdot(\mathbf{v}\times\mathbf{l}) = \mathbf{r}\cdot(GM\mathbf{u}+\mathbf{C}) &= \mathbf{r}\cdot\left(GM\frac{\mathbf{r}}{r}\right) + r\mathbf{u}\cdot\left((r_0v_0^2-GM)\mathbf{i}\right)\\ &= GMr + r(r_0v_0^2-GM)\cos\theta\,.\end{aligned} \tag{100}$$

Comparing these two equations, we see

$$r_0^2v_0^2 = GMr + r(r_0v_0^2-GM)\cos\theta\,, \tag{101}$$

or

$$r = \frac{\frac{r_0^2v_0^2}{GM}}{1+\left(\frac{r_0v_0^2}{GM}-1\right)\cos\theta} = \frac{k}{1+e\cos\theta}\,, \tag{102}$$

where

$$k = \frac{r_0^2v_0^2}{GM}\,,\quad e = \frac{r_0v_0^2}{GM}-1\,. \tag{103}$$

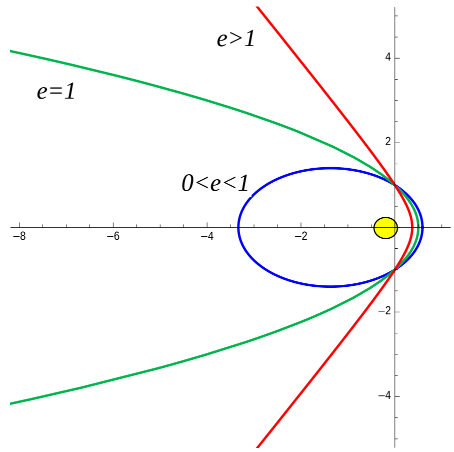

Figure 17: Planetary orbits as conic sections.

In other words, this is the equation of a conic section with $k = de$ and k and e given above. This can result in an ellipse ($0 < e < 1$), parabola ($e = 1$) or hyperbola ($e > 1$). The conic that the orbit sketches is the result of the mass of the body causing the gravitational force, as well as the initial position and velocity. If $e \geq 1$, the orbit is not closed, i.e. the bodies escapes the gravitational pull. The condition $e = 1$ give the escape velocity,

$$v_{\rm esc} = \sqrt{\frac{2GM}{r_0}}\,. \tag{104}$$

2.1.11.2 Deriving Kepler's Laws

For the first law, since we know that the planets don't escape, the orbit must be an ellipse by the previous discussion. The fact that the sun is at a focus comes from equation (102). Since r is the distance to a focus and the distance to the centre of the gravitational force, in the case our the planets in our solar system, the Sun must be at the ellipse. To find the second law, we equate the two expressions for $\mathbf{l}$ (86) and (94). This gives us

$$r^2 \frac{d\theta}{dt} = r_0 v_0\,. \tag{105}$$

Since the curve should be described by a function of the angle, $r = f(\theta)$, and hence form the formula for area in polar coordinates,

$$A = \int_{\theta_1}^{\theta_2} \frac{1}{2}[f(\theta)]^2 \, d\theta \,, \tag{106}$$

we can find the rate of change of the area

$$\frac{dA}{dt} = \frac{dA}{d\theta}\frac{d\theta}{dt} = \frac{1}{2}[f(\theta)]^2\frac{d\theta}{dt} = \frac{1}{2}r^2\frac{d\theta}{dt} = \frac{1}{2}r_0 v_0 \,, \tag{107}$$

which is a constant. Since the area changes at a constant rate, the areas traced out in equal times are equal.

For the third law, we begin with the area of an ellipse

$$\pi ab \,, \tag{108}$$

where a and b are the semi-major and semi-minor axes. Although we haven't actually covered this is is quite easy to prove (try!). In one period, the area the radial line sweeps out is

$$\int_0^T \frac{dA}{dt}\, dt = \int_0^T \frac{1}{2} r_0 v_0 \, dt = \frac{1}{2} r_0 v_0 T \,, \tag{109}$$

and therefore

$$\frac{1}{2} r_0 v_0 T = \pi ab \,. \tag{110}$$

We can square this to obtain

$$T^2 = \frac{4\pi^2 a^2 b^2}{r_0^2 v_0^2} \,. \tag{111}$$

We know, however, that $c^2 = a^2 - b^2$ for an ellipse and also

$$e = \frac{c}{a} = \frac{\sqrt{a^2 - b^2}}{a} \,, \tag{112}$$

which implies

$$b^2 = a^2(1 - e^2) \,. \tag{113}$$

Moreover, we can see from equation (103) that

$$\frac{r_0^2 v_0^2}{GM} = k = a(1 - e^2)\,, \tag{114}$$

and so

$$T^2 = \frac{4\pi^2 a^3}{r_0^2 v_0^2} k = \frac{4\pi^2 a^3}{r_0^2 v_0^2} \frac{r_0^2 v_0^2}{GM} = \frac{4\pi^2}{GM} a^3\,, \tag{115}$$

and taking the square root gives the result,

$$T = \frac{2\pi}{\sqrt{GM}} a^{3/2}\,. \tag{116}$$