1 Introduction

We want to to describe curves in space, and then possibly motion along these curves or even motion of curves themselves. Intuitively curves are some sort of 'one-dimensional' sets of points in space.

Consider first a curve in the plane. One approach to describing a curve would be to consider the graph of a function $f(x)$. The graph is the set of points (x, y) where $y = f(x)$. For the graph to be what we think of as a curve, we need f to be continuous and the domain of f to be an interval of real numbers. There are problems with this approach. For example, it can be awkward even for relatively simple curves, for example a circle, and it is *unnatural to treat x and y differently.*

To motivate a better description and the definition of curves we shall actually use, look at the roller coaster. This brings us to vector functions and parametrisations.

1.1 Vector functions

You know what a vector is. (A basic review is given at the end of the chapter.) The position of a point can be represented as a vector (**its position vector**) from the origin to the point. We typically use the vector $\mathbf{r}$ to denote position and write

$$\mathbf{r} = (x, y) \in \mathbb{R}^2$$

for a point in the plane and

$$\mathbf{r} = (x, y, z) \in \mathbb{R}^3$$

for a point in space.
So x, y, and possibly z are the components of the position vector. These are also the usual Cartesian coordinates for a point. At times we consider the general case where the dimension is some unspecified n. In these cases we label the components of position by $\mathbf{r} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

You know what a real valued function of a real variable is:

$$f : U \subseteq \mathbb{R} \to \mathbb{R}.$$

To each real number in U, a subset of $\mathbb{R}$, f assigns a value in $\mathbb{R}$.

Definition 1. *A vector function*

$$\mathbf{r} : U \subseteq \mathbb{R} \to \mathbb{R}^n$$

assigns a point in $\mathbb{R}^n$, that is a vector, to each real number in U.

The independent variable is usually denoted by t, from "time". Frequently we call $\mathbf{r}$ a map and think of it as taking points in $\mathbb{R}$ and mapping them to points in $\mathbb{R}^n$. We sometime write $t \mapsto \mathbf{r}(t)$.
Vector function can be viewed in terms of component functions

$$\boxed{\mathbf{r}(t) = (x_1(t), x_2(t), \cdots, x_j(t), \cdots, x_n(t)),}$$

where each component function $x_j(t)$ is a real valued function. You can think of a vector function as just a collection of real-valued functions arranged as components of a vector.

We are primarily interested in either the plane ($n = 2$) or the space ($n = 3$), and in these cases we usually use the notation $x(t), y(t)$, and possibly $z(t)$ for component functions rather than $x_1(t), x_2(t)$, and $x_3(t)$. For example, in space we would typically write

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

Things you know about vectors and functions can now be applied to vector functions in the obvious way, e.g. given vector functions

$$\mathbf{f}(t) = f_1(t)\mathbf{i} + f_2(t)\mathbf{j}, \quad \mathbf{g}(t) = g_1(t)\mathbf{i} + g_2(t)\mathbf{j},$$

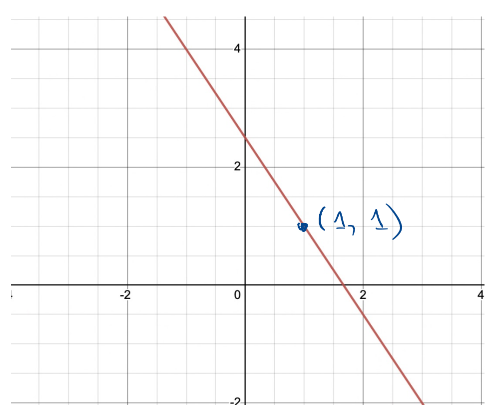
and scalar α, we could define the vector function $\mathbf{r}$

$$\begin{aligned} \mathbf{r}(t) &= \mathbf{f}(t) + \alpha\mathbf{g}(t) \\ &= (f_1(t) + \alpha g_1(t))\,\mathbf{i} + (f_2(t) + \alpha g_2(t))\,\mathbf{j} \end{aligned}$$

Example 1. *Come up with a vector function whose values trace a single straight line.*

$\underline{r} : \mathbb{R} \longrightarrow \mathbb{R}^2,$

$\underline{r}(t) = (1 + 2 \cdot t,\ 1 - 3 \cdot t)$



2 Curves. Definitions

There is a close connection between vector functions and curves. Essentially a curve is the set of points traced out by a vector function as the independent variable ranges over an interval $I = [a, b]$.

Definition 2. *Let* $\mathbf{r} : I \to \mathbb{R}^n$ *be a continuous vector function, where* $I \subseteq \mathbb{R}$ *is an interval. Then the set of points*

$$\mathcal{C} = \{\mathbf{r} \mid t \in I\}$$

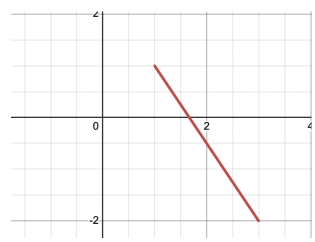
is a curve. The mapping $\mathbf{r}$ *is called a parametrisation of the curve* $\mathcal{C}$*.*

Remarks.

- The definitions requires $\mathbf{r}$ to be a continuous vector function, $\lim_{t\to t_0} \mathbf{r}(t) = \mathbf{r}(t_0)$ for any $t_0 \in I$ (more on this next week). The continuity means what you think it does: it guarantees that there are no breaks in the curve $\mathcal{C}$.

- We will generically write the interval as $I = [a, b]$, but it is understood that I could be infinite or semi-infinite. For example: $I = (-\infty, \infty), I = (-\infty, 0]$, or $I = [1, \infty)$. Example:

  $\underline{r}: [0,1] \to \mathbb{R}^2$,
  $\underline{r}(t) = (1 + 2\cdot t, 1 - 3\cdot t)$



- You should be aware that there is some variation in the definition of a curve. In some texts a curve is defined as the map $\mathbf{r}$ not the set of points $\mathcal{C}$. Some authors require I to be a closed interval, others do not.

- There are infinitely many possible parametrisations for a given curve $\mathcal{C}$. (Example: If $\mathbf{r} : I \to \mathbb{R}^n$ is a parametrisation of $\mathcal{C}$ and $h : I \to I$ is onto and continuous, then $\mathbf{r} \circ h$ is also a parametrisation of $\mathcal{C}$.) Two parametrisations are called *equivalent* if they define the same curve. Equivalence classes of parametrisations are curves in the sense of our definition. Examples:

  $\underline{r}_1 : \mathbb{R} \to \mathbb{R}^2$,
  $\underline{r}_1(t) = (1 + 2\cdot t, 1 - 3\cdot t)$

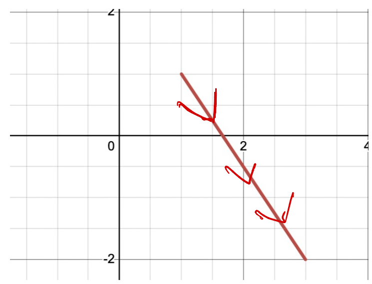
  and $\underline{r}_2 : \mathbb{R} \to \mathbb{R}$
  $\underline{r}_2(t) = (1-t)\cdot(1, 1) + t\cdot(11; -14)$

  give the same curve.

- The definition in terms of parametrisations gives us something in addition to the set of points $\mathcal{C}$: if $\mathbf{r} : I \to \mathcal{C}$ is injective, it defines an orientation the direction of “travel” along $\mathcal{C}$ corresponding to increasing t. There are two possible orientations as we can travel either from $\mathbf{r}(a)$ to $\mathbf{r}(b)$ or in the opposite direction. (Injectivity is sufficient, but not necessary for existence of orientation.)

$\underline{r} : [0,1] \longrightarrow \mathbb{R}^2,$

$\underline{r}(t) = (1 + 2\cdot t,\ 1 - 3\cdot t)$



- If all we remember about $\mathbf{r}$ is its image, we refer to $\mathcal{C}$ simply as a curve; if we remember $\mathrm{Im}(\mathbf{r})$ and and the orientation, then $\mathcal{C}$ is called an oriented curve; if we remember the map $\mathbf{r}$ itself, then we are talking about a parametrised curve.

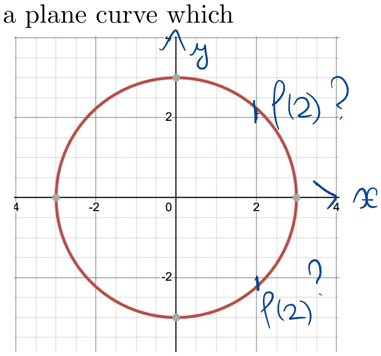
- From the way we motivated the subject, one sees a close connection between curves (purely geometrical objects) and particle paths (trajectories associated with motion in time). This is further reinforced by using t for the independent variable.

  It is generally not necessary to distinguish these cases. However, if a curve actually arises from the motion of a ”particle”, we shall say path rather than curve. When the independent parameter is actual time, then there are specific physical meanings to expressions defined later.

  Particle paths may be such that same points in space are retraced multiple times. If we are only interested in the curve itself, we would probably not consider a parametrisation covering the same points multiple times, while if we are thinking of a particle path, then we probably would.

- Graphs of continuous functions are curves: the graph $\{(x, f(x)) \in \mathbb{R}^2 \mid x \in [a, b]\} = \mathrm{Im}(\mathbf{r})$, where $\mathbf{r} : [a, b] \to \mathbb{R}^2$ is such that $\mathbf{r}(x) = (x, f(x))$.

Question. Can you come up with an example of a plane curve which is not the graph of a function?

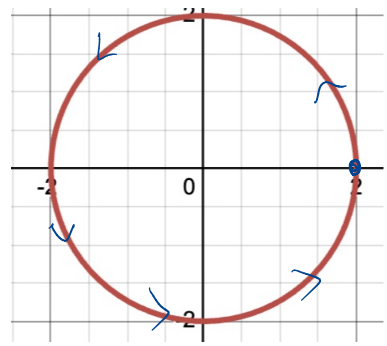


What is $f(2)$?

2.1 Working with parameterisations

Examples:

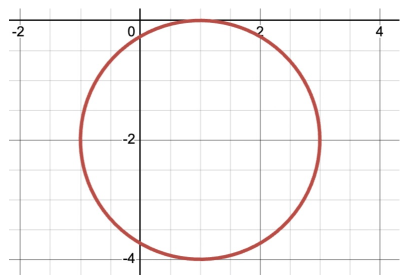
$$\mathbf{r}(t) = (R\cos t, R\sin t), t \in [0, 2\pi]$$



$R = 2$

Fix $a, b \in \mathbb{R}$. Then consider

$$\mathbf{r}(t) = (a + R\cos t, b + R\sin t), t \in [0, 2\pi]$$



$a = 1$
$b = -2$
$R = 2$

Now, if b, c are fixed, what does

$$\mathbf{r}(t) = (0, b + R\cos t, c + R\sin t), t \in [0, 2\pi]$$

parameterise?

If $\mathbf{r}(t) = (x(t), y(t))$ is a parametrisation for some curve, then $(ax(t), y(t))$ will stretch the curve in x direction if $a > 1$ and compress the curve in x direction if $0 < a < 1$. For example

$$\mathbf{r}(t) = (a \cos t, b \sin t), t \in [0, 2\pi]$$

is the parametrisation for an ellipse if $a \neq b$.

Given $a \in \mathbb{R}$, what does

$$\mathbf{r}(t) = (t, at^2), t \in \mathbb{R}$$

parameterise?. How about $\mathbf{r}(t) = (t^3, at^6)$, $t \in \mathbb{R}$? The same question for $\mathbf{r}(t) = (\sin t, a \sin^2 t)$.

- $\mathbf{r}(t) = (b\sinh(t), a\cosh(t)), t \in \mathbb{R}$ and a, b positive constants, parametrises the top branch of the hyperbola $y^2/a^2 - x^2/b^2 = 1$, see the review material for more information on conic sections.

- $\mathbf{r}(t) = (\text{ at } \cos t, \text{ at } \sin t), t \geq 0$ is the parametrisation of an Archimedean spiral, (a spiral given in polar coordinates by $r = a\theta$). $\mathbf{r}(t) = (e^{at}\cos t, e^{at}\sin t)$, $t \in \mathbb{R}$ is the parametrisation of a Logarithmic spiral, (in polar coordinates $r = e^{a\theta}$).

  Some curves can be given in polar coordinates (r, θ), where $r = f(\theta)$ is a function of the angle. Such curves are called **polar graphs** or **graphs of polar functions**. The Logarithmic spiral defined above is easily expressed as a polar function.

- Speaking of polar coordinates, one can also consider parametrised curves in polar coordinates. For example, what do you think is the curve parameterised by

$$(r(t), \theta(t)) = (2R\sin t, t), t \in [0, \pi]?$$

  This parametrisation can be re-written as a polar function $r = f(\theta) = 2R\sin(\theta)$.

- In space: $\mathbf{r}(t) = (R\cos t, R\sin t, kt), t \in \mathbb{R}, k \neq 0$, is the parametrisation of a helix.

- In space: $\mathbf{r}(t) = \left(e^t \cos bt, e^t \sin bt, e^t\right), t \in \mathbb{R}\ b \neq 0$, is the parametrisation of a curve winding around a cone (circular conical surface).

3 Sketching

Using a computer software, curve sketching is relatively easy, so you should learn to use Matlab, Python (Matplotlib), Desmos (2D plotting) or Geogebra 3D calculator to sketch curves. Here is some guidance for situations in which you don't have a calculator.

- Spot rotation in the form of cos and sin terms in two of the coordinates. For example, if $x(t) = b(t)\cos(t)$ and $y(t) = b(t)\sin(t)$, then in polar coordinates $\theta = t$ and $r = b(\theta)$.

- Spot powers. Look for eliminating t between two coordinates. For example, getting $y = x^\alpha$ or $x = y^\alpha$, which you should know how to sketch.

- In the above cases it might be desirable to first shift coordinates.

- Look at extreme value of t, i.e. $t \to \pm\infty$, and other particular value of t such as $t = 0$.

- Space curves can be quite challenging. Look to separate behaviour in the different coordinate directions. Choose a good orientation and draw the axes first. Good luck.

4 Finding parameterisations

Finding a parametrisation is somewhat the opposite of sketching. Given some description or specification of a curve, you need to produce a parametrisation for it, generally for the purpose of doing some further calculation involving the curve. This can be quite challenging, but we will stick to some easy examples.

It is good to be able to parameterise conic sections (see the Review section). Then, another good skill is to be able to parameterise curves made of straight segments. Then, it's good to be able to extend these to combinations and cases where circle radius varies, e.g. a spiral. It is useful to be prepared to think in polar coordinates and to switch between polar and Cartesian coordinates as necessary.

4.1 Curves formed by multiple segments

We carefully defined curves and parametrisations so that curves are continuous images of a single interval I. Yet, when is comes to many applications this can be quite annoying and unnecessary. Consider parametrising the plane curve $\mathcal{C}$ which starts at $(1,0)$ and proceed to $(-1,0)$ along the upper unit semicircle and then returns to $(1,0)$ along the x-axis. The annoyance is that curve is easily described as the union of two segments, each of which is easily parametrised:

$$\begin{aligned} \mathbf{r}_1(t) &= \cos(t)\mathbf{i} + \sin(t)\mathbf{j}, \qquad & t \in [0,\pi] \\ \mathbf{r}_2(t) &= t\mathbf{i}, & t \in [-1,1] \end{aligned}$$

Of course one can adjust things to produce a mapping over a single interval I, (probably you would adjust the second parametrisation). However, for most things we care about this is simply unnecessary. It is sufficient to work with a set of parametrisations, $\mathbf{r}_1(t), \cdots, \mathbf{r}_k(t)$ defined on intervals $I_1(t), \cdots, I_k(t)$, if these naturally describe a curve composed of a union of k segments.

4.2 Further concepts

There are a few further concepts related to curves which are worth defining, even though these will not be essential in our discussion.

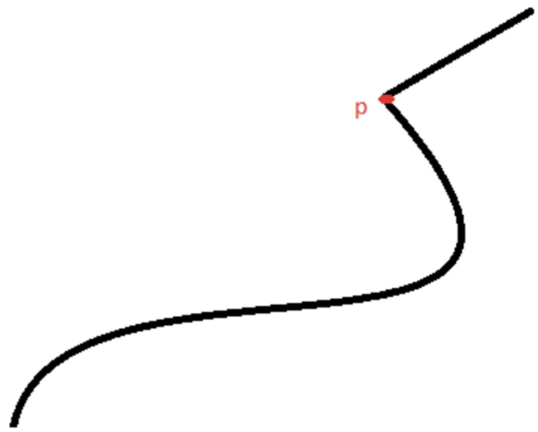
A curve $\mathcal{C}$ is closed if $I = [a,b] \to \mathbb{R}^2$ and $\mathbf{r}(a) = \mathbf{r}(b)$. A closed curve is also called a loop.

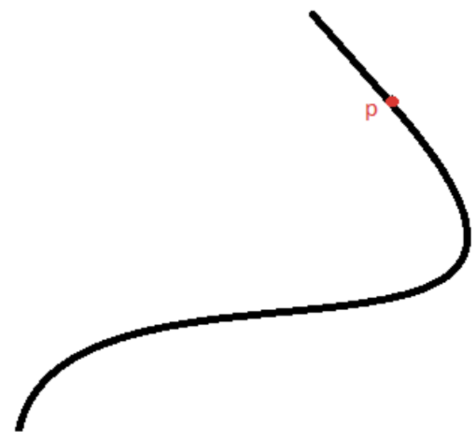
Remark. Be careful, some very 'non-loop-like' curves turn out to be closed under the above definition, for example, $\mathbf{r}(t) = (\sin(t), 0), t \in [0, 2\pi]$. The corresponding curve is the interval $\{(x,0) || x | \leq 1\}$ in $\mathbb{R}^2$. It is also unsatisfactory to define the geometric notion of closedness for a parametrised curve, rather than a curve. To circumvent these difficulties, it turns out to be easier to define closedness for **embedded curves** (see below) first and generalise from there, see `https://en.wikipedia.org/wiki/Curve` for details and references.

Definition 3. *A curve is embedded if it does not self intersect. In terms of parametrisations, a curve is embedded if it can be parametrised by a mapping $\mathbf{r} : I \to \mathbb{R}^n$ that is injective except possibly at the end points of $I = [a,b]$. We need to allow the possibility that $\mathbf{r}(a) = \mathbf{r}(b)$ since this just means that the curve is closed and does not correspond to an intersection.*

From the definition of injective, a curve is embedded if it can be parametrised by $\mathbf{r}(t)$ where $t_1 \neq t_2$ implies $\mathbf{r}(t_1) \neq \mathbf{r}(t_2)$, except if $t_1 = a$ and $t_2 = b$.

Definition 4. *A curve is regular if there exits a parametrisation such that its derivative $\mathbf{r}'$ is defined and nonzero at all points. We have not yet defined the derivative $\mathbf{r}'$, but we will next week. The important point is that regular curves have no corners or cusps and hence are nice to deal with. We will say more next week.*





5 Review

5.1 Real vectors

This is just a quick reminder of some basic things you already be familiar with.

Notation. Vectors will be denoted by boldface and sometime over arrow, e.g. $\mathbf{v}$ or $\vec{v}$. In lectures, underline will be used instead of bold face, e.g. $\underline{v}$. We only consider real vectors in this course, so $\mathbf{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$. The real numbers $v_1, v_2, \ldots, v_n$ are call the components of $\mathbf{v}$.
The term **scalar** reffers to a real number, to contrast with vectors.

The dot product. You should know properties of the dot product and how to compute it. Recall that

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta$$

where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$, or as

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} u_i v_i,$$

where u_i, v_i are the components of u and v. The dot product of two vectors is a scalar, and $u \cdot v = v \cdot u$. Notice that $||v|| = (v \cdot v)^{1/2}$.

You should be familiar with important unit vectors. For example, the basis vectors for Cartesian coordinates in three dimensions are

$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \mathbf{k} = (0, 0, 1)$$

Then we can write the position vector $\mathbf{r}$ as a sum of components times basis vectors:

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

One sometimes uses the following notation for these bases vectors:

$$\mathbf{e}_1 = (1, 0, 0) \quad \mathbf{e}_2 = (0, 1, 0) \quad \mathbf{e}_3 = (0, 0, 1)$$

Given a nonzero vector $\mathbf{v}$, that is $\mathbf{v} \neq (0, 0, \ldots, 0) = \mathbf{0}$, the direction of $\mathbf{v}$ is: $\hat{\mathbf{v}} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$. So each nonzero vector can be written

$$\mathbf{v} = \|\mathbf{v}\| \hat{\mathbf{v}} = (\text{ magnitude }) \times (\text{ direction }) = (\text{ scalar }) \times (\text{ unit vector })$$

5.2 Polar coordinates

You should be familiar with polar coordinates, typically denoted (r, θ). The relationship between polar and Cartesian coordinate is:

$$\begin{aligned} x &= r \cos \theta, \quad r^2 = x^2 + y^2 \\ y &= r \sin \theta, \quad \tan \theta = \tfrac{y}{x} \end{aligned}$$

You should know how to go back and forth between Cartesian and polar coordinates and be able to use polar coordinates as needed. There is a small issue here with non-uniqueness of polar coordinates. You should know that coordinates $(r, \theta), (r, \theta + 2\pi)$, and $(-r, \theta + \pi)$ all correspond to the same point in the plane. One sometimes works with $r \geq 0$ and θ restricted to a fixed range to avoid the non-uniqueness. In this case I generally prefer to work with $\theta \in [0, 2\pi)$, but this depends on the problem. Sometimes it is more convenient to work with $\theta \in (-\pi, \pi]$. Sometimes it is best to live with the non-uniqueness. For $r = 0$ the value of θ is irrelevant.

5.3 Conic sections

The following curves are conic sections: circle, ellipse, parabola, and hyperbola. These curves all come from intersecting a circular conical surface with a plane, hence the name conic section. The curves also come about as sets of points (x, y) satisfying the general quadratic equation:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

where A, B, C, D, E, F are constants with not all three of A, B, and C are equal to zero. These curves appear frequently and you should know standard forms for the equations in each case:

- **Circle**:

  Points satisfying $x^2 + y^2 = a^2$.

- **Ellipse:**

  Points satisfying $\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1$

- **Hyperbola:**

  Points satisfying $\left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2 = 1$

- **Parabola**: Points satisfying $y = ax^2$, or sometimes for consistency with the other expressions $y^2 = 4ax$. (In the first case the parabola opens upward or downward depending on the sign of a, while in the second case it opens to the left or right.)

6 Problems for the seminar

1. Sketch the curves parameterised by

   (a) $\mathbf{r}(t) = (t^4, t^2), t \in \mathbb{R}$;

   (b) $\mathbf{r}(t) = (t^5, t^3), t \in \mathbb{R}$;

   (c) $\mathbf{r}(t) = t(\sin t, \cos t),\ t \in \mathbb{R}$;

   (d) $\mathbf{r}(t) = (1, 3\cos t, \sin t), t \in [0, 2\pi)$.

   use an arrow to indicate the direction in which t increases. In most cases, you cannot show the whole curve. Just show a representative portion.

2. Sketch the curves parameterised in polar coordinates by

$$(r, \theta) = (\sin t, t), \quad t \in [0, \pi] \quad (r, \theta) = \left(e^{t/2\pi}, t\right), \quad t \in [-2\pi, 2\pi]$$

Note that each of these parametrisations can be written as $r = f(\theta)$. Identify the geometric shape of the curve from the first example.

3. Parameterise the circle in the plane whose centre is $(-1, -1)$ and radius is $\sqrt{2}$ in polar coordinates.

4. Give a parametrisation of the ellipse that is centred on the point $(2, 2)$, has semi-major axis from the centre to the point $(4, 2)$, and has semi-minor axis from the centre to the point $(2, 3)$.

5. Imagine going on a bicycle ride through the country. The tires stay in contact with the road and rotate in a predictable pattern. Now suppose a very determined ant is tired after a long day and wants to get home. So it hangs onto the side of the tire and gets a free ride. The path that this ant travels down a straight road is called a **cycloid**. Find parameterisations for a cycloid generated by a bycicle wheel which has the shape of a circle with radius radius R.

P1. Sketch the curves parameterised by

(a) $\mathbf{r}(t) = \left(t^2, t^3\right), \quad t \in \mathbb{R};$

(b) $\mathbf{r}(t) = t\cos t\mathbf{i} + |t|\sin t\mathbf{j}, \quad t \in \mathbb{R};$

(c) $\mathbf{r}(t) = \frac{1+e^{-t}}{1+2e^{-t}}(\cos t, -\sin t), \quad t \in \mathbb{R};$

(d) $\mathbf{r}(t) = \cos(t)\mathbf{i} + \sin(t)\mathbf{j} + t^2\mathbf{k}, \quad t \in \mathbb{R}.$

In cases where you cannot show the whole curve, just show a representative portion. Use arrows to indicate the direction in which t increases. These sketches must be accompanied by brief explanations (and a few words/calculations) justifying their shapes; For each curve state, with justification, whether or not it is embedded and whether or not it is closed.

P2. (a) Sketch the plane curve parameterised by

$$\mathbf{r}(\tau) = (2 + \cos(3\tau))\cos(2\tau)\mathbf{i} + (2 + \cos(3\tau))\sin(2\tau)\mathbf{j}, \quad \tau \in [0, 2\pi].$$

Find the points of self-intersection.

(b) The curve in part (a) can be viewed as a two-dimensional projection of a three-dimensional embedded curve called the *trefoil knot:*

$$\mathbf{r}(\tau) = (2 + \cos(3\tau))\cos(2\tau)\mathbf{i} + (2 + \cos(3\tau))\sin(2\tau)\mathbf{j} + \sin(3\tau)\mathbf{k}, \quad \tau \in [0, 2\pi].$$

Show explicitly that this curve does not self-intersect. Reproduce the sketch from part (a), namely the $2D$ sketch, with only x and y coordinates, indicating with gaps where the $3D$ curve passes over itself, as it would be observed from the positive part of the z-axis.