ADVANCED CALCULUS

MA1132

HILARY TERM 2017/2018

DR. ANTHONY BROWN

Contents

1. Vector-Valued Functions

In the majority of the maths you will have met so far, the codomain of the functions you will have been dealing with will usually have been the set of real numbers $\mathbb{R}$, although you will also have met mappings in linear algebra, where the codomain will be sets of matrices, and perhaps you have also met complex functions where the codomain will have been the set of complex numbers $\mathbb{C}$.

In this chapter we will investigate the properties of functions that have as their codomain vector spaces. We will restrict our attention to $\mathbb{R}^2$ and $\mathbb{R}^3$, but in general the codomain of a function can be any vector space, even an infinite dimensional one!

1.1. **Introduction to Vector-Valued Functions.**

Let us start with a definition.

Definition 1.1.1. A *Vector-Valued Function* is a function of the form

$$\begin{aligned}\mathbf{r}\colon A &\to B\\ t &\mapsto \mathbf{r}(t),\end{aligned}$$

where A is any set and B is a vector space (or possibly a subset of a vector space).

Remark 1.1.2. Vectors and vector-valued functions are denoted by bold type in typed maths, but when hand written, they should be denoted in some other way. For example, they can be underlined, or written with an arrow across the top.

While in the most general situation A can be any set and B can be any vector space (or a subset of any vector space), in this chapter we will restrict our attention to the cases where A will either be $\mathbb{R}$ or a subset of $\mathbb{R}$ and B will either be $\mathbb{R}^2$ or $\mathbb{R}^3$. So in this chapter a typical vector-valued function will be of the form

$$\begin{aligned} \mathbf{r}\colon \mathbb{R} &\to \mathbb{R}^2 \\ t &\mapsto \mathbf{r}(t), \end{aligned}$$

or

$$\begin{aligned} \mathbf{r}\colon [0,1] &\to \mathbb{R}^3 \\ t &\mapsto \mathbf{r}(t), \end{aligned}$$

or something similar.

While we can write the image of a real number t as $\mathbf{r}(t)$, it will often be more convenient to write it in terms of the *component functions*, so we will write $\mathbf{r}(t) = (x(t), y(t))$ (when the codomain is $\mathbb{R}^2$) or $\mathbf{r}(t) = (x(t), y(t), z(t))$ (when the codomain is $\mathbb{R}^3$), where x, y and z are functions that map $\mathbb{R}$ (or a subset of $\mathbb{R}$) to $\mathbb{R}$.

Sometimes the domain of a vector-valued function will not be given, and in these cases we take the domain to be the largest one that 'makes sense'. This will be the intersection of the domains of the component functions.

Next we want to consider how to graphically represent vector-valued functions. Of course in the general case where the codomain is any (possibly infinite dimensional) vector space this is not a very profitable direction to consider. However when the codomain is $\mathbb{R}^2$ or $\mathbb{R}^3$, as will be the case in this chapter, a graphical approach can be very useful. Let us start by giving the formal definition of graph of a function.

Definition 1.1.3. Given a function

$$\begin{aligned} f\colon A &\to B \\ x &\mapsto f(x), \end{aligned}$$

then its *graph* is defined to be the set $\{(x, f(x))\colon x \in A\}$.

Thus the graph of a function is defined to be a **set** whose elements are ordered pairs. When we are dealing with familiar functions that map $\mathbb{R}$ to $\mathbb{R}$, then each element in each ordered pair is a number, so we can represent each ordered pair as a point in $\mathbb{R}^2$. The graph of such a function (if it is sufficiently 'well behaved') is then obtained by joining up these points.

The approach we take with vector-valued functions is slightly different in that the 'graph' that we end up with (in two or three dimensions as appropriate) is not a direct representation of the formal graph of the function, since we only trace the image of the function, and the domain of the function is not represented.

For example, if we want to plot the 'graph' of the function

$$\begin{aligned} \mathbf{r}\colon \mathbb{R} &\to \mathbb{R}^2 \\ t &\mapsto \mathbf{r}(t) = (x(t), y(t)), \end{aligned}$$

then we plot the set of points $\{(x(t), y(t))\colon t \in \mathbb{R}\}$. Similarly if we want to plot the 'graph' of the function

$$\begin{aligned} \mathbf{r}\colon [0,1] &\to \mathbb{R}^3 \\ t &\mapsto \mathbf{r}(t) = (x(t), y(t), z(t)), \end{aligned}$$

then we plot the set of points $\{(x(t), y(t), z(t))\colon t \in [0,1]\}$, although we would more usually use a program to plot this, unless we happen to be very good at drawing lines in three dimensions.

Remark 1.1.4.

(a) Sometimes we will want to indicate the direction a point on the graph moves as we increase t, and we will do this using an arrow on the graph.
(b) It might also be useful in certain circumstances to remind ourselves that the points we are plotting are vectors rather than just points, and in this case we can draw the *radius vector* (or *position vector*) from the origin to a given point on the graph.

1.2. **Calculus of Vector-Valued Functions.**

Having introduced vector-valued functions in Section 1.1, in this section we will attempt to generalize some results from the calculus you are familiar with to this more general setting.

Since the continuity and derivative of a function $f\colon \mathbb{R} \to \mathbb{R}$ is defined in term of a limit, to make any progress towards generalizing these concepts for vector-valued functions, we first have to define what a limit means for such functions.

The definition we give below will not work for any vector space since it uses the norm of a vector, and the formal definition of a vector space makes no mention of a distance or norm. It will work for all normed vector spaces though, even those that are infinite dimensional (provided the domain is $\mathbb{R}$ or a subset of $\mathbb{R}$).

Definition 1.2.1. Let $a, b, c \in \mathbb{R}$ be such that $b < a < c$. Then given a vector-valued function $\mathbf{r}$ whose domain contains a punctured open interval $(b, c)\backslash\{a\}$, we say that

$$\lim_{t\to a} \mathbf{r}(t) = \mathbf{l} \quad \text{if} \quad \lim_{t\to a} \|\mathbf{r}(t) - \mathbf{l}\| = 0.$$

Remark 1.2.2.

(a) As we would expect, the limit of a vector-valued function is a vector.

(b) Given a vector valued function $\mathbf{r}$ and a vector $\mathbf{l}$, then the limit $\lim_{t\to a} \|\mathbf{r}(t) - \mathbf{l}\|$ is a limit of a real function f defined by

$$\begin{aligned} f\colon (b,c)\backslash\{a\} &\to \mathbb{R} \\ t &\mapsto \|\mathbf{r}(t) - \mathbf{l}\|, \end{aligned}$$

so we know how to calculate this limit from our knowledge of real functions.

(c) The definition we have given only works for vector-valued functions that have domain $\mathbb{R}$ or a subset of $\mathbb{R}$. The definition can be further generalized to situations where the domain is also a normed space. This creates a problem in that we do not yet know how to calculate the limit $\lim_{\mathbf{t}\to\mathbf{a}} \|\mathbf{r}(\mathbf{t}) - \mathbf{l}\|$ in these cases. In Chapter 2, we will look at such limits for functions with codomain $\mathbb{R}$, but in this course we won't look at the most general case, where a function maps a vector space to another vector space.

It is often the case in mathematics that the definition of a certain property does not get us very far in terms of being able to calculate things. This is the case here, but the following theorem enables us to easily calculate limits, at least for vector-valued functions that have $\mathbb{R}^2$ or $\mathbb{R}^3$ as their codomain.

Theorem 1.2.3. *Let $a, b, c \in \mathbb{R}$ be such that $b < a < c$, and let the function $\mathbf{r}$ be defined by*

$$\begin{aligned} \mathbf{r}\colon (b,c)\backslash\{a\} &\to \mathbb{R}^2 \\ t &\mapsto \mathbf{r}(t) = (x(t), y(t)), \end{aligned}$$

then the limit $\lim_{t\to a} \mathbf{r}(t)$ exists if an only if both the limits $\lim_{t\to a} x(t)$ and $\lim_{t\to a} y(t)$ exist. Moreover, if the limits exist,

$$\lim_{t\to a} \mathbf{r}(t) = \Big(\lim_{t\to a} x(t), \lim_{t\to a} y(t)\Big) = \lim_{t\to a} x(t)\mathbf{i} + \lim_{t\to a} y(t)\mathbf{j}.$$

Remark 1.2.4.

(a) This theorem can be extended to vector-valued functions with codomain $\mathbb{R}^3$ (or even $\mathbb{R}^n$) in the obvious manner.

(b) It is also true that if $\lim_{t\to a} \mathbf{r}(t)$ exists then $\lim_{t\to a} x(t)$ equals the first component of $\lim_{t\to a} \mathbf{r}(t)$ and $\lim_{t\to a} y(t)$ equals the second component of $\lim_{t\to a} \mathbf{r}(t)$, but we won't usually use the theorem in this form.

Now that we know how to calculate limits of vector-valued functions, it is easy to define when they are continuous, by essentially using the same definition as for real-valued functions.

Definition 1.2.5. A vector-valued function $\mathbf{r}$ is said to be *continuous* at $t = a$ if

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

Remark 1.2.6. It follows from this definition that a vector-valued function is continuous at a point (respectively interval) if and only if its component functions are continuous at this point (respectively interval).

The definition of differentiability also uses essentially the same definition as for real-valued functions.

Definition 1.2.7. A vector-valued function $\mathbf{r}$ defined on an open interval containing t is said to be *differentiable* at t if the limit

$$\lim_{h\to 0} \frac{\mathbf{r}(t+h)-\mathbf{r}(t)}{h}$$

exists. If this is the case, then we define the *derivative* of $\mathbf{r}$ at t to be

$$\mathbf{r}'(t) = \lim_{h\to 0} \frac{\mathbf{r}(t+h)-\mathbf{r}(t)}{h}.$$

Remark 1.2.8.

(a) As we would expect, the function $\mathbf{r}'$ is a vector-valued function.
(b) As is the case with real-valued functions, there are different notations used instead of $\mathbf{r}'$. For example $\dfrac{d}{dt}(\mathbf{r})$, $\dfrac{d\mathbf{r}}{dt}$ or $\dot{\mathbf{r}}$

In the case of real-valued functions, the derivative at a point has a geometrical interpretation as the slope of the tangent to the graph of the function at the point. The derivative of a vector-valued function also has a very useful geometrical interpretation: If a vector-valued function $\mathbf{r}$ is differentiable at a point t with non-zero derivative, then the direction of the vector $\mathbf{r}'(t)$ gives the direction of the tangent to the graph of $\mathbf{r}$ at the point $\mathbf{r}(t)$, pointing in the direction of increasing t. Of course this also makes it easy to find the equation of the tangent line to the curve at a point $\mathbf{r}(t)$ (provided $\mathbf{r}'(t) \neq \mathbf{0}$); it is $\mathbf{r}(t) + p\mathbf{r}'(t)$, $p \in \mathbb{R}$.

Again, similarly to the case of real-valued functions, we don't usually use Definition 1.2.7 to calculate derivatives. For this we use the following theorem.

Theorem 1.2.9. *Let $a, b, c \in \mathbb{R}$ be such that $b < a < c$, and let the function $\mathbf{r}$ be defined by*

$$\begin{aligned} \mathbf{r}\colon (b,c) &\to \mathbb{R}^2 \\ t &\mapsto \mathbf{r}(t) = (x(t), y(t)), \end{aligned}$$

then $\mathbf{r}$ is differentiable at $t = a$ if an only if both x and y are differentiable at $t = a$. Moreover, if both x and y are differentiable at $t = a$, then

$$\mathbf{r}'(a) = (x'(a), y'(a)) = x'(a)\mathbf{i} + y'(a)\mathbf{j}.$$

Remark 1.2.10. This theorem can be extended to vector-valued functions with codomain $\mathbb{R}^3$ (or even $\mathbb{R}^n$) in the obvious manner.

There are also various rules of differentiation we can use to help us differentiate vector valued functions.

Theorem 1.2.11. *Suppose that* $\mathbf{r}$ *and* $\mathbf{s}$ *are vector-valued functions differentiable at a point* t*, that* f *is a real-valued function differentiable at* t*, and that* $k \in \mathbb{R}$ *is a constant. Then*

(a) $(k\mathbf{r})'(t) = k\mathbf{r}'(t)$.
(b) $(\mathbf{r} + \mathbf{s})'(t) = \mathbf{r}'(t) + \mathbf{s}'(t)$.
(c) $(f\mathbf{r})'(t) = f(t)\mathbf{r}'(t) + f'(t)\mathbf{r}(t)$.

Although Theorem 1.2.11(c) is a form of the product rule, we cannot have a proper product or quotient rule simply because multiplication and division of vectors is not defined. However there are two way of combining vectors that have some similarities with multiplication, the dot and cross products, and the following theorem tells us how to differentiate vector-valued functions that have been combined using these.

Theorem 1.2.12. *Suppose that* $\mathbf{r}$ *and* $\mathbf{s}$ *are vector-valued functions differentiable at a point* t*. Then*

(a) $(\mathbf{r} \cdot \mathbf{s})'(t) = \mathbf{r}(t) \cdot \mathbf{s}'(t) + \mathbf{r}'(t) \cdot \mathbf{s}(t)$.
(b) $(\mathbf{r} \times \mathbf{s})'(t) = \mathbf{r}(t) \times \mathbf{s}'(t) + \mathbf{r}'(t) \times \mathbf{s}(t)$, *provided the codomain of* $\mathbf{r}$ *and* $\mathbf{s}$ *is* $\mathbb{R}^3$.

Remark 1.2.13.

(a) Since the dot product of two vectors yields a number and the cross product of two vectors yields a vector, it is to be expected that the derivative of a dot product at a point t is a number, while the derivative of a cross product at a point t is a vector.
(b) The rule for dot products works provided $\mathbf{r}$ and $\mathbf{s}$ have the same dimension n for their codomain (and is generalizable to infinite dimensional spaces, where it is usually called an *inner product*), but the rule for cross products only works if this dimension is three. This is because the cross product is only defined for three-dimensional vectors.

Having has a look at differentiation, we will now proceed to look at integration. As with real-valued functions, this can be defined in terms of Riemann Sums, but in this course we will concentrate of the mechanics of integration, rather than the theory. As was the case with differentiation, the method of integrating vector-valued functions is to work with the integrals of the component functions.

Theorem 1.2.14. *Let the function*

$$\begin{aligned}\mathbf{r}\colon [a,b] &\to \mathbb{R}^2\\ t &\mapsto \mathbf{r}(t) = (x(t), y(t))\end{aligned}$$

be continuous. Then

$$\int_a^b \mathbf{r}(t)\,dt = \left(\int_a^b x(t)\,dt\right)\mathbf{i} + \left(\int_a^b y(t)\,dt\right)\mathbf{j}.$$

Remark 1.2.15.

(a) As we would expect, when we integrate a vector-valued function, we end up with a vector.
(b) Of course this theorem can be extended in the obvious way to vector-valued functions with codomain $\mathbb{R}^n$.

The multiple and sum rules also carry over to vector-valued integration.

Theorem 1.2.16. *Let the functions*

$$\begin{aligned} \mathbf{r}\colon [a,b] &\to \mathbb{R}^2 \\ t &\mapsto \mathbf{r}(t) \end{aligned}$$

and

$$\begin{aligned} \mathbf{s}\colon [a,b] &\to \mathbb{R}^2 \\ t &\mapsto \mathbf{s}(t) \end{aligned}$$

be continuous, and let $k \in \mathbb{R}$ be a constant. Then

(a) $\displaystyle\int_a^b k\mathbf{r}(t)\,dt = k\int_a^b \mathbf{r}(t)\,dt.$
(b) $\displaystyle\int_a^b \mathbf{r}(t) + \mathbf{s}(t)\,dt = \int_a^b \mathbf{r}(t)\,dt + \int_a^b \mathbf{s}(t)\,dt.$

Remark 1.2.17. Of course Theorem 1.2.16 can be extended in the obvious way to vector-valued functions with codomain $\mathbb{R}^n$.

The concepts of antiderivative and indefinite integral also generalize to vector-valued functions.

Definition 1.2.18.

(a) An *antiderivative* of a vector-valued function $\mathbf{r}$ is any vector-valued function $\mathbf{R}$ such that $\mathbf{R}'(t) = \mathbf{r}(t)$.
(b) The *indefinite integral* of a vector-valued function $\mathbf{r}$ is defined to be

$$\int \mathbf{r}(t)\,dt = \mathbf{R}(t) + \mathbf{c},$$

where $\mathbf{R}$ is any antiderivative of $\mathbf{r}$, and $\mathbf{c}$ is a (general) constant vector.

Remark 1.2.19.

(a) As we would expect, the indefinite integral of a vector-valued function is a vector-valued function.
(b) If $\mathbf{r}(t) = (x(t), y(t))$ and if X is an antiderivative of x and Y is an antiderivative of y, then it follows from the above definition and Theorem 1.2.9 that $(X(t), Y(t))$ is an antiderivative of $\mathbf{r}$, and similarly for vector-valued functions with codomain $\mathbb{R}^n$.

(c) The familiar concept that integration and differentiation are inverse operations carry over to this more general setting. In particular

$$\left(\int \mathbf{r}(t)\,dt\right)' = \mathbf{r}(t) \quad \text{and} \quad \int \mathbf{r}'(t)\,dt = \mathbf{r}(t) + \mathbf{c}.$$

(d) The multiple and sum rules also hold for indefinite integrals.

We will finish this chapter with the following more general version of the Fundamental Theorem of Calculus which follows from Theorem 1.2.14, Remark 1.2.19(b) and the real-valued version of the Fundamental Theorem of Calculus.

Theorem 1.2.20. *Let* $\mathbf{R}$ *be an antiderivative of* $\mathbf{r}$ *on the interval* $[a, b]$. *Then*

$$\int_a^b \mathbf{r}(t)\,dt = [\mathbf{R}(t)]_a^b = \mathbf{R}(b) - \mathbf{R}(a).$$

1.3. **Change of Parameter; Arc Length.**

When dealing with real-valued functions, the graph is unique to each function. That is, if we have a different function then we will have a different graph. However, recall from Section 1.1 that what we are calling the 'graph' of a vector-valued function is not formally its graph. So we can have different vector-valued functions that yield the same 'graph'. Some of these functions will be easier to calculate with than others, and in fact there is one specific one, the so called *arc length* or *unit speed* parametrization that is particularly useful.

In order to find this special parametrization, we first have to calculate the arc length of a section of the graph of a vector-valued function, and this is what we will now do.

It follows immediately from the formula used to find the length of a vector, that if $\mathbf{r}(t) = (x(t), y(t))$ is a differentiable vector-valued function, then

$$\|\mathbf{r}'(t)\| = \sqrt{x'(t)^2 + y'(t)^2},$$

with a similar formula if $\mathbf{r}(t) = (x(t), y(t), z(t))$.

Thus we can write the formula for calculating the arc length L of a section of a parametrized curve between the points $\mathbf{r}(a)$ and $\mathbf{r}(b)$ as

$$L = \int_a^b \|\mathbf{r}'(t)\|\,dt.$$

In view of Theorem 1.2.14, in order to perform this integration, we need the function $\|\mathbf{r}'(t)\|$ to be continuous, which is true if and only if the function $\mathbf{r}'(t)$ is continuous. While we could perform the integration if $\mathbf{r}'(t) = \mathbf{0}$, this can lead to corners on our graphs, which we will not be dealing with in this course. Thus we make the following definition of the types of parametrizations that we will be dealing with.

Definition 1.3.1. A curve represented by the graph of a vector-valued function $\mathbf{r}$ is said to be *smoothly parametrized* by $\mathbf{r}$ if $\mathbf{r}'(t)$ is continuous and non-zero for all t in the domain of $\mathbf{r}$. If this is the case, then we say that $\mathbf{r}(t)$ is a *smooth* function.

Remark 1.3.2. This usage of the word smooth does not tally with its usual usage, where it means a function that can be differentiated any number of times.

Given any curve smoothly parametrized by $\mathbf{r}$, we now want to see how to find another parametrization which will be a unit speed parametrization. Let us start with a definition.

Definition 1.3.3. An *arc length* or *unit speed* parametrization is a parametrization where the parameter measures the distance along the curve.

Put more mathematically, this means that if $\mathbf{r}$ is an arc length parametrization, then

$$\int_a^b \|\mathbf{r}'(s)\|\, ds = b - a.$$

Remark 1.3.4.

(a) We will use $\mathbf{r}(s)$ rather than $\mathbf{r}(t)$ when we want to denote a vector-valued function that is a unit speed parametrization (s for speed).
(b) We will see later on where the name unit speed comes from.

Now, given any smooth parametrization $\mathbf{r}(t)$, we want to find a unit speed smooth parametrization $\mathbf{r}(s)$ which traces out the same curve. In order to make sure the new parametrization is smooth, we will need the following generalization of the chain rule

Theorem 1.3.5. *Let $\mathbf{r}(t)$ be a differentiable vector-valued function. If $t = g(s)$, where g is differentiable with respect to s then $\mathbf{r}(s)$ is also differentiable with respect to s and*

$$\frac{d\mathbf{r}}{ds} = \frac{d\mathbf{r}}{dt} \cdot \frac{dt}{ds}. \tag{1}$$

Remark 1.3.6. Since $\dfrac{d\mathbf{r}}{dt}$ is a vector and $\dfrac{dt}{ds}$ is a number, it would be more usual to write $\dfrac{dt}{ds} \cdot \dfrac{d\mathbf{r}}{dt}$. However this would make this version of the chain rule look different to the version for real-values functions, so I think it is best to stick to $\dfrac{d\mathbf{r}}{ds} = \dfrac{d\mathbf{r}}{dt} \cdot \dfrac{dt}{ds}$.

Looking at Equation (1), we see that if we want $\mathbf{r}$ to be a smooth function of s then all we need is for $\dfrac{dt}{ds}$ to be continuous and non-zero (at least for the range of s we are interested in), since we will be starting with $\mathbf{r}$ as a smooth function of t, so $\dfrac{d\mathbf{r}}{dt}$ will be continuous and non-zero.

Remark 1.3.7. Since we want $\dfrac{dt}{ds}$ to be continuous and non-zero for the range of s we are interested in, it follows that $\dfrac{dt}{ds}$ must either be always positive or always negative for this range of s. If it is always positive, then we say we have a positive change of parameter and if it is always negative, then we say we have a negative change of parameter. Geometrically, a positive change of parameter preserves the direction of the path, while a negative change of parameter reverses it.

We are now in a position to describe how to find the function $t = g(s)$ which will enable us to convert a smooth parametrization $\mathbf{r}(t)$ into a smooth unit speed parametrization $\mathbf{r}(s)$. Given a smooth parametrization $\mathbf{r}(t)$ and a given number t_0, let us define

$$s(t) = \int_{t_0}^{t} \|\mathbf{r}'(v)\|\, dv. \tag{2}$$

Remark 1.3.8.

(a) We have to introduce the 'dummy variable' v to integrate against here since we are using t as a limit of integration.
(b) If we differentiate both sides of this equation, then it follows from the Fundamental Theorem of Calculus (real-valued version) that $s'(t) = \|\mathbf{r}'(t)\|$.

Warning 1.3.9. This is **NOT** the function g that we want. Here we are expressing s as a function of t, so to get the function g that we want (that is t as a function of s), we have to find the inverse of this function.

Once we have this inverse function, to find the unit speed parametrization $\mathbf{r}(s)$, it is a simple matter of writing t in terms of s in $\mathbf{r}(t)$.

To conclude this section, let us see why a unit speed parametrization is so called. If we have obtained a unit speed parametrization using Equation (2), as we noted in Remark 1.3.8(b), $s'(t) = \|\mathbf{r}'(t)\|$. Written in Leibniz notation, that is $\dfrac{ds}{dt} = \left\|\dfrac{d\mathbf{r}}{dt}\right\|$. Hence, using Equation (1), the rule for differentiating inverse functions and the fact that $\dfrac{ds}{dt}$ is positive, it follows that

$$\left\|\frac{d\mathbf{r}}{ds}\right\| = \left\|\frac{d\mathbf{r}}{dt}\cdot\frac{dt}{ds}\right\| = \left\|\frac{d\mathbf{r}/dt}{ds/dt}\right\| = \frac{\|d\mathbf{r}/dt\|}{\|ds/dt\|} = \frac{\|d\mathbf{r}/dt\|}{ds/dt} = 1. \tag{3}$$

Now, one interpretation of $\left\|\dfrac{d\mathbf{r}}{ds}\right\|$ is the 'speed' a point moves along the curve, and this is why unit speed parametrizations are so called.

1.4. **Unit Tangent, Normal and Binormal Vectors.**

This section and the next will consist of a series of definitions of vectors and constants that are derived from a given parametrization of a curve. It can be quite hard to visualize what is going on, so as you are reading through these sections, it can be useful to keep one eye on the animation that is about a quarter the way down the page at https://en.wikipedia.org/wiki/Frenet-serret formulas. This link will take you directly to the animation.

In Section 1.2, we already noted that the derivative $\mathbf{r}'(t)$ gives a vector in the direction of the tangent to the graph of $\mathbf{r}$ at $\mathbf{r}(t)$ in the direction of increasing t. Thus we can easily find a unit vector in this direction.

Definition 1.4.1. If $\mathbf{r}$ is a smooth vector-valued function with codomain $\mathbb{R}^2$ or $\mathbb{R}^3$, then the *unit tangent vector* to the graph of $\mathbf{r}$ at $\mathbf{r}(t)$ is given by

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}. \tag{4}$$

Remark 1.4.2.

(a) Since $\mathbf{T}(t)$ is a unit tangent vector to a curve, then at a particular point $\mathbf{r}(t)$, it only depends on the curve, not on the parametrization (although of course, for a different parametrization, the value of t at a given point on the curve may be different).
(b) In Equation (3) we showed that if $\mathbf{s}$ is a unit speed parametrization then $\|\mathbf{r}'(s)\| = 1$. So in this case Equation (4) reduces to $\mathbf{T}(s) = \mathbf{r}'(s)$.

Using Theorem 1.2.12(a), we have

$$\frac{d}{dt}\left[\|\mathbf{T}(t)\|^2\right] = \frac{d}{dt}\left[\mathbf{T}(t)\cdot\mathbf{T}(t)\right] = \mathbf{T}(t)\cdot\mathbf{T}'(t) + \mathbf{T}'(t)\cdot\mathbf{T}(t) = 2\mathbf{T}(t)\cdot\mathbf{T}'(t).$$

However $\|\mathbf{T}(t)\|^2 = 1^2 = 1$, so that $\dfrac{d}{dt}\left[\|\mathbf{T}(t)\|^2\right] = \dfrac{d}{dt}(1) = 0$. Thus we also have $\mathbf{T}(t)\cdot\mathbf{T}'(t) = 0$. In other words if $\mathbf{T}'(t) \neq \mathbf{0}$ then it is orthogonal to $\mathbf{T}(t)$ ($\mathbf{T}(t)$ has norm 1 so it can't be the zero vector). This leads to the following definition.

Definition 1.4.3. If $\mathbf{r}$ is a smooth vector-valued function with codomain $\mathbb{R}^2$ or $\mathbb{R}^3$ and if $\|\mathbf{T}'(t)\| \neq 0$, then the *principal normal vector* to the graph of $\mathbf{r}$ at $\mathbf{r}(t)$ is given by

$$\mathbf{N}(t) = \frac{\mathbf{T}'(t)}{\|\mathbf{T}'(t)\|}. \tag{5}$$

Remark 1.4.4.

(a) Sometimes we drop the word 'principal' and just call $\mathbf{N}(t)$ the normal vector to the graph of $\mathbf{r}$ at $\mathbf{r}(t)$. However note that in two dimensions there are two directions for a general normal vector, and there are in three dimensions, there are infinitely many directions.
(b) Since the principal normal vector is defined in terms of the rate of change of the tangent vector, it points in the direction towards which the path is curving.
(c) The condition $\|\mathbf{T}'(t)\| \neq 0$ means that straight lines do not have a principal normal vector defined (there is no change in the direction of the tangent vector in this case since it always points in the direction of the straight line). Of course there are vectors normal to the line, but there is no principal one.

(d) Since $\mathbf{T}(s) = \mathbf{r}'(s)$ if $\mathbf{r}(s)$ is a unit speed parametrization, it follows that in this case, Equation (5) reduces to

$$\mathbf{N}(s) = \frac{\mathbf{r}''(s)}{\|\mathbf{r}''(s)\|}.$$

We noted above that in three dimensions, there are infinitely many possible normal vectors to $\mathbf{T}(t)$. As well as $\mathbf{N}(t)$, there is another particular normal vector that we define. This is the vector that is normal to both $\mathbf{T}(t)$ and $\mathbf{N}(t)$.

Definition 1.4.5. If $\mathbf{r}$ is a smooth vector-valued function with codomain $\mathbb{R}^3$ and if $\|\mathbf{T}'(t)\| \neq 0$, then the *binormal vector* to the graph of $\mathbf{r}$ at $\mathbf{r}(t)$ is given by

$$\mathbf{B}(t) = \mathbf{T}(t) \times \mathbf{N}(t). \tag{6}$$

Remark 1.4.6.

(a) The binormal vector is not defined for $\mathbb{R}^2$ since we can't move out of the plane containing $\mathbf{T}$ and $\mathbf{N}$ in $\mathbb{R}^2$.
(b) Since $\mathbf{T}(t)$ and $\mathbf{N}(t)$ are orthogonal and are unit vectors, it follows from Equation (6) that $\mathbf{B}(t)$ is also a unit vector.
(c) At each point $\mathbf{r}(t)$, the triple $\{\mathbf{T}(t), \mathbf{N}(t), \mathbf{B}(t)\}$ forms a right-handed orthonormal set, so they can be used as an orthonormal basis for a coordinate system. This is sometimes called the **TNB** frame, or the Frenet-Serret frame. It follows directly from this that

$$\mathbf{T}(t) \times \mathbf{N}(t) = \mathbf{B}(t), \quad \mathbf{N}(t) \times \mathbf{B}(t) = \mathbf{T}(t), \quad \text{and} \quad \mathbf{B}(t) \times \mathbf{T}(t) = \mathbf{N}(t).$$

Similarly

$$\mathbf{N}(t) \times \mathbf{T}(t) = -\mathbf{B}(t), \quad \mathbf{B}(t) \times \mathbf{N}(t) = -\mathbf{T}(t), \quad \text{and} \quad \mathbf{T}(t) \times \mathbf{B}(t) = -\mathbf{N}(t).$$

(d) We also sometimes give names to the planes containing two of the vectors $\mathbf{T}(t)$, $\mathbf{N}(t)$ and $\mathbf{B}(t)$. The **TB** frame is called the rectifying frame, the **TN** frame is called the osculating plane and the **NB** frame is called the normal plane.
(e) Using Equation (4), Equation (5) and Equation (6), we can also express $\mathbf{B}(t)$ in terms of $\mathbf{r}(t)$ as

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

and in the case of a unit speed parametrization, this simplifies to

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

1.5. **Curvature.**

In Section 1.5, we defined various vectors associated with the graph of a vector-valued function. In this section, we will turn our attention to other properties such a graph may have. Since this section is entitled 'Curvature', you may think that we only have one property to deal with, but this is not the case.

In two dimensions, it is the case since we can measure the curvature of a curve at any point with just one number, but in three dimensions things are more complicated since there can be curvature in the **N** direction but the **TN** plane can also tilt as we move along the curve.

Here is the definition of one of these curvatures.

Definition 1.5.1. If $\mathbf{r}$ is a smooth vector-valued function with codomain $\mathbb{R}^2$ or $\mathbb{R}^3$ and if s is a unit speed parametrization, then the *curvature* of the graph of $\mathbf{r}$ at $\mathbf{r}(s)$ is given by

$$\kappa(s) = \|\mathbf{T}'(s)\|. \tag{7}$$

Remark 1.5.2.

(a) The curvature is defined to be the norm of a vector, so it is a number rather than a vector. In particular it does not have a direction.
(b) This measure of curvature measures the curvature in the **TN** frame. If the curvature is constant and the curve does not leave the **TN** frame, then the curve will be a circle. The radius of this circle will be the reciprocal of the curvature.
(c) For a unit speed parametrization, we have $\mathbf{T}(s) = \mathbf{r}'(s)$, so we can also write Equation (7) as $\kappa(s) = \|\mathbf{r}''(s)\|$.
(d) Rearranging Equation (5), we obtain $\mathbf{T}'(s) = \|\mathbf{T}'(s)\|\mathbf{N}(s)$, so it follows from Equation (7) that $\mathbf{T}'(s) = \kappa(s)\mathbf{N}(s)$.

As is often the case in mathematics, the definition is not usually the best way of calculating the curvature, since to do this, we would first need to find a unit speed parametrization. A couple of easier methods are given in the following theorem.

Theorem 1.5.3. *If* $\mathbf{r}$ *is a smooth vector-valued function with codomain* $\mathbb{R}^2$ *or* $\mathbb{R}^3$ *and if* $\mathbf{T}'(t)$ *and* $\mathbf{r}''(t)$ *exist, with* $\mathbf{r}'(t)$ *being non-zero, then*

(a)

$$\kappa(t) = \frac{\|\mathbf{T}'(t)\|}{\|\mathbf{r}'(t)\|}.$$

(b)

$$\kappa(t) = \frac{\|\mathbf{r}'(t) \times \mathbf{r}''(t)\|}{\|\mathbf{r}'(t)\|^3}.$$

Proof.

(a) Using Remark 1.3.8(b),

$$\begin{aligned}
\kappa(t) &= \|\mathbf{T}'(s)\| \\
&= \left\| \frac{d\mathbf{T}}{ds} \right\| \\
&= \left\| \frac{d\mathbf{T}}{dt} \cdot \frac{dt}{ds} \right\| \\
&= \left\| \frac{d\mathbf{T}/dt}{ds/dt} \right\| \\
&= \frac{\|\mathbf{T}'(t)\|}{s'(t)} \\
&= \frac{\|\mathbf{T}'(t)\|}{\|\mathbf{r}'(t)\|} \\
&= \frac{\|\mathbf{T}'(t)\|}{\|\mathbf{r}'(t)\|}.
\end{aligned}$$

(b) By Equation (4),

$$\mathbf{r}'(t) = \|\mathbf{r}'(t)\|\mathbf{T}(t). \tag{8}$$

Then, on differentiating Equation (8) using Theorem 1.2.11(c), we obtain

$$\begin{aligned}
\mathbf{r}''(t) &= \|\mathbf{r}'(t)\|'\mathbf{T}(t) + \|\mathbf{r}(t)\|\mathbf{T}'(t) \\
&= \|\mathbf{r}'(t)\|'\mathbf{T}(t) + \|\mathbf{r}'(t)\| \cdot \|\mathbf{T}'(t)\|\mathbf{N}(t) \quad \text{using (5)} \\
&= \|\mathbf{r}'(t)\|'\mathbf{T}(t) + \kappa(t)\|\mathbf{r}'(t)\|^2\mathbf{N}(t) \quad \text{by Part (a).} \qquad (9)
\end{aligned}$$

We then take the cross product of Equation (8) and Equation (9), and note that the cross product of parallel vectors is $\mathbf{0}$, to get

$$\begin{aligned}
\mathbf{r}'(t) \times \mathbf{r}''(t) &= \|\mathbf{r}'(t)\|\mathbf{T}(t) \times \|\mathbf{r}'(t)\|'\mathbf{T}(t) + \|\mathbf{r}'(t)\|\mathbf{T}(t) \times \kappa(t)\|\mathbf{r}'(t)\|^2\mathbf{N}(t) \\
&= \kappa(t)\|\mathbf{r}'(t)\|^3\mathbf{T}(t) \times \mathbf{N}(t) \\
&= \kappa(t)\|\mathbf{r}'(t)\|^3\mathbf{B}(t).
\end{aligned}$$

If we take the norm of both sides of this equation then, since $\kappa(t)$ is a non-negative number and $\mathbf{B}(t)$ is a unit vector, we get

$$\|\mathbf{r}'(t) \times \mathbf{r}''(t)\| = \kappa(t)\|\mathbf{r}'(t)\|^3\|\mathbf{B}(t)\| = \kappa(t)\|\mathbf{r}'(t)\|^3,$$

and the result follows.

□

Remark 1.5.4. If we are working in $\mathbb{R}^2$ and if $\mathbf{r}'(t) = (x'(t), y'(t))$, then we have to write $\mathbf{r}'(t) = (x'(t), y'(t), 0)$ and similarly for $\mathbf{r}''(t)$ in order to take the cross product in Part (b).

We now turn to the other measure of curvature.

Definition 1.5.5. If $\mathbf{s}$ is a smooth vector-valued function with codomain $\mathbb{R}^3$, then the *torsion* of the graph of $\mathbf{r}$ at $\mathbf{r}(s)$ is given by

$$\tau(s) = -\mathbf{N}(s) \cdot \mathbf{B}'(s). \tag{10}$$

Remark 1.5.6.

(a) We have used the parameter s implying that we are working with a unit speed parametrization.
(b) As with the curvature κ, $\tau(s)$ is a number rather than a vector.
(c) This measure of curvature measures how quickly the $\mathbf{TN}$ frame rotates as we move along the curve.
(d) If $\tau(s) = 0$, then the $\mathbf{TN}$ frame doesn't rotate, so the whole curve lies in a plane.
(e) It is not defined for $\mathbb{R}^2$ since in $\mathbb{R}^2$, the $\mathbf{TN}$ frame can't rotate, since we only have two dimensions to play with.
(f) If we are dealing with a parametrization which is not unit speed, then we need to use the equation

$$\tau(t) = \frac{\mathbf{r}'(t) \cdot (\mathbf{r}''(t) \times \mathbf{r}'''(t))}{\|\mathbf{r}'(t) \times \mathbf{r}''(t)\|^2}.$$

There are various formulae involving κ and τ; the most important of these are called the *Frenet-Serret* formulae. We have already mentioned the first of these in Remark 1.5.2(d), but we will collect them together in the next theorem for reference, but we will omit the proofs of the second and third.

Theorem 1.5.7 (The Frenet-Serret Formulae)**.**

(a) $\mathbf{T}'(s) = \kappa(s)\mathbf{N}(s)$.
(b) $\mathbf{N}'(s) = -\kappa(s)\mathbf{T}(s) + \tau(s)\mathbf{B}(s)$.
(c) $\mathbf{B}'(s) = -\tau(s)\mathbf{N}(s)$.

Remark 1.5.8.

(a) We have used the parameter s implying that we are working with a unit speed parametrization.
(b) We can also combine the three formulae into one matrix formula:

$$\begin{bmatrix} \mathbf{T}' \\ \mathbf{N}' \\ \mathbf{B}' \end{bmatrix} = \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix} \begin{bmatrix} \mathbf{T} \\ \mathbf{N} \\ \mathbf{B} \end{bmatrix}.$$

(c) For a non-unit speed parametrization, the corresponding matrix equation is

$$\begin{bmatrix} \mathbf{T}' \\ \mathbf{N}' \\ \mathbf{B}' \end{bmatrix} = \begin{bmatrix} 0 & \kappa\dfrac{ds}{dt} & 0 \\ -\kappa\dfrac{ds}{dt} & 0 & \tau\dfrac{ds}{dt} \\ 0 & -\tau\dfrac{ds}{dt} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{T} \\ \mathbf{N} \\ \mathbf{B} \end{bmatrix} = \frac{ds}{dt} \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix} \begin{bmatrix} \mathbf{T} \\ \mathbf{N} \\ \mathbf{B} \end{bmatrix}.$$

1.6. **Motion Along a Curve.**

In the first five sections, we have introduced vector-valued functions, studied their calculus and used this calculus to analyse various properties of their graphs. In this section, we put this accumulated knowledge to study motion along a curve.

The key is that instead of regarding the graph as a whole, we think of t as representing time, and look at how the graph is traced out as t changes. In particular, we take $\mathbf{r}(t)$ to represent the position of a particle at time t. Since if we differentiate position, we get velocity and if we differentiate velocity, we get acceleration, we have the following.

Proposition 1.6.1. *If* $\mathbf{r}$ *is a twice differentiable vector-valued function and if the position of a particle at time* t *is given by* $\mathbf{r}(t)$*, then*

(a) The velocity of the particle at time t*,* $\mathbf{v}(t)$*, is given by* $\mathbf{r}'(t)$*.*
(b) The acceleration of the particle at time t*,* $\mathbf{a}(t)$*, is given by* $\mathbf{r}''(t)$*.*

Remark 1.6.2. If we represent the arc length by s then the speed of the particle is given by

$$\|\mathbf{v}\| = \|\mathbf{r}'\| = \left\|\frac{d\mathbf{r}}{dt}\right\| = \left\|\frac{d\mathbf{r}}{ds}\cdot\frac{ds}{dt}\right\| = \left\|\frac{d\mathbf{r}}{ds}\right\|\cdot\left|\frac{ds}{dt}\right| = \left|\frac{ds}{dt}\right| = \frac{ds}{dt}, \tag{11}$$

the second last equality following from Equation (3) and the last equality following from the fact that s increases as t increases, so $\dfrac{ds}{dt}$ is positive.

Since, by definition, $\mathbf{T}(t) = \dfrac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}$, we also have $\mathbf{r}'(t) = \|\mathbf{r}'(t)\|\mathbf{T}(t)$ (provided $\mathbf{r}'(t) \neq \mathbf{0}$) . Hence it follows from Equation (11), that we can write

$$\mathbf{v}(t) = \mathbf{r}'(t) = \frac{ds}{dt}(t)\mathbf{T}(t). \tag{12}$$

That is, the velocity of the particle is equal to the speed of the particle, $\dfrac{ds}{dt}(t)$, times a unit vector in the direction of the velocity of the particle, $\mathbf{T}(t)$.

As you would expect, we can also go in the reverse direction, and integrate the acceleration to get the velocity, and integrate the velocity to get the position. Of course if we use indefinite integration then a (vector) constant of integration will appear. Perhaps the two most useful things we can find are:

- The displacement of the particle from time t_1 to time t_2, which is given by
$$\int_{t_1}^{t_2} \mathbf{v}(t)\,dt = [\mathbf{r}(t)]_{t_1}^{t_2} = \mathbf{r}(t_2) - \mathbf{r}(t_1).$$
- The total distance travelled of the particle from time t_1 to time t_2, which is given by
$$\int_{t_1}^{t_2} \|\mathbf{v}(t)\|\,dt.$$

It can also be useful to resolve the acceleration into that in the direction of travel and that at right angles to the direction of travel. Or put another way, into the $\mathbf{T}$ direction and into the $\mathbf{N}$ direction.

Proposition 1.6.3. *If* $\mathbf{r}$ *is a twice differentiable vector-valued function and if the position of a particle at time* t *is given by* $\mathbf{r}(t)$*, then (provided* $\mathbf{r}'(t) \neq \mathbf{0}$*) its acceleration is given by*

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

where s*,* $\mathbf{T}$*,* $\mathbf{N}$ *and* κ *have their usual meanings.*

Proof. Using Equation (12) and the fact that acceleration is the derivative of velocity

$$\begin{aligned}
\mathbf{a} &= \frac{d}{dt}\left(\frac{ds}{dt}\mathbf{T}\right) \\
&= \frac{d^2s}{dt^2}\mathbf{T} + \frac{ds}{dt}\cdot\frac{d\mathbf{T}}{dt} \\
&= \frac{d^2s}{dt^2}\mathbf{T} + \frac{ds}{dt}\cdot\frac{d\mathbf{T}}{ds}\cdot\frac{ds}{dt} \\
&= \frac{d^2s}{dt^2}\mathbf{T} + \left(\frac{ds}{dt}\right)^2\frac{d\mathbf{T}}{ds} \\
&= \frac{d^2s}{dt^2}\mathbf{T} + \left(\frac{ds}{dt}\right)^2\kappa\mathbf{N} \\
&= \frac{d^2s}{dt^2}\mathbf{T} + \kappa\left(\frac{ds}{dt}\right)^2\mathbf{N},
\end{aligned}$$

where the second last line follows from Remark 1.5.2(d). $\square$

Remark 1.6.4. The scalar $\dfrac{d^2s}{dt^2}$ is called the *tangential scalar component of acceleration* and is denoted a_T. The scalar $\kappa\left(\dfrac{ds}{dt}\right)^2$ is called the *normal scalar component of acceleration* and is denoted a_N. Thus we can write

$$\mathbf{a} = a_T\mathbf{T} + a_N\mathbf{N}. \tag{13}$$

Similarly $\dfrac{d^2s}{dt^2}\mathbf{T}$ is called the *tangential vector component of acceleration* and $\kappa\left(\dfrac{ds}{dt}\right)^2\mathbf{N}$ is called the *normal vector component of acceleration*.

The following theorem gives alternative methods for calculating a_T, a_N and κ.

Theorem 1.6.5. *If* $\mathbf{r}$ *is a twice differentiable vector-valued function and if the position of a particle at time* t *is given by* $\mathbf{r}$*, the velocity by* $\mathbf{v}$ *and the acceleration by* $\mathbf{a}$*, then provided* $\|\mathbf{v}\| \neq \mathbf{0}$*,*

(a) The tangential scalar component of acceleration is given by $a_T = \dfrac{\mathbf{v}\cdot\mathbf{a}}{\|\mathbf{v}\|}$*.*

(b) The tangential scalar component of acceleration is given by $a_N = \dfrac{\|\mathbf{v}\times\mathbf{a}\|}{\|\mathbf{v}\|}$*.*

(c) The curvature is given by $\kappa = \dfrac{\|\mathbf{v}\times\mathbf{a}\|}{\|\mathbf{v}\|^3}$*.*

Proof.

(a) Since $\mathbf{v}$ and $\mathbf{T}$ are in the same direction and since $\mathbf{T}$ has length 1, their dot product is $\|\mathbf{v}\|$. Also, since $\mathbf{v}$ and $\mathbf{N}$ are at right angles their dot product is zero. Thus we have

$$\mathbf{v}\cdot\mathbf{a} = \mathbf{v}\cdot(a_T\mathbf{T} + a_N\mathbf{N}) = a_T\mathbf{v}\cdot\mathbf{T} + a_N\mathbf{v}\cdot\mathbf{N} = a_T\|\mathbf{v}\|,$$

and the result follows on dividing by $\|\mathbf{v}\|$.

(b) Since $\mathbf{v}$ and $\mathbf{T}$ are in the same direction, their cross product is zero. Also, since $\mathbf{v}$ and $\mathbf{N}$ are at right angles and since $\mathbf{N}$ has length 1, their cross product is $\mathbf{v}$. Thus we have

$$\begin{aligned}\|\mathbf{v}\times\mathbf{a}\| &= \|\mathbf{v}\times(a_T\mathbf{T} + a_N\mathbf{N})\,\| \\ &= \|a_T\mathbf{v}\times\mathbf{T} + a_N\mathbf{v}\times\mathbf{N}\| \\ &= |a_N|\cdot\|\mathbf{v}\| \\ &= a_N\|\mathbf{v}\|,\end{aligned}$$

the last step following since a_N is non-negative. The result then follows on dividing by $\|\mathbf{v}\|$.

(c) By Equation (11), $\dfrac{ds}{dt} = \|\mathbf{v}\|$, so this follows from Part (b) and the definition of a_N on dividing the equation $\kappa\left(\dfrac{ds}{dt}\right)^2 = a_N = \dfrac{\|\mathbf{v}\times\mathbf{a}\|}{\|\mathbf{v}\|}$ by $\left(\dfrac{ds}{dt}\right)^2 = \|\mathbf{v}\|^2$.

□

Remark 1.6.6.

(a) Since $\mathbf{T}$ and $\mathbf{N}$, are orthogonal and of length 1, it follows from Equation (13) and Pythagoras, that $\|\mathbf{a}\|^2 = a_T^2 + a_N^2$. So, if $\|\mathbf{a}\|$ is known, then it is usually easier to find a_N using Theorem 1.6.5(a) and $a_N = \sqrt{\|\mathbf{a}\|^2 - a_T^2}$, since in general, cross products can be hard to calculate.

(b) Comparing (b) and (c), we see that $\kappa = \dfrac{a_N}{\|\mathbf{v}\|^2}$.