III Differentiation

15 Differentiation as a Limit 55
15.1 The Limit of a Function
15.2 The Derivative

16 Derivatives of Basic Functions 61

17 Derivatives of Compound Functions . . 63
17.1 Derivative of a Sum/Difference
17.2 The Product Rule
17.3 The Quotient Rule
17.4 The Chain Rule
17.5 Inverse Differentiation

18 Stationary Points and Extrema 71

19 Implicit and Parametric Differentiation 73
19.1 Implicit Differentiation
19.2 Parametric Differentiation

20 Vector Differentiation 77

15. Differentiation as a Limit (Riley 2.1.1)

Differentiation will be our first subject in the field of **Calculus** (lat. "Small Pebble"). It concerns the study of functions and specifically how they change locally, i.e. **infinitesimally** close to a point. Differentiation will require the notion of an infinitesimal limit. While we discuss this in more detail in Chapter 6 on "Series and Limits", we here require first a basic definition of a limit in order to understand differentiation.

In this chapter we restrict the discussion mostly to functions $f = f(x)$ of one variable x.

15.1 The Limit of a Function

Definition 15.1 — Limit of a Function. Consider a function $f(x)$. If we can make $f(x)$ as close as we want to a given number L by choosing the variable x sufficiently close to a number a, then L is said to be the limit of $f(x)$ as $x \to a$ and it is written as

$$\lim_{x\to a} f(x) = L. \tag{15.1}$$

We illustrate this in a couple of examples.

■ **Example 15.1** In the case the function is defined at the point $x = a$, the limit is simply given by the function value at that point, e.g.

$$\lim_{x\to 2} x^2 = 4.$$

■

■ **Example 15.2** On the other hand, a limit can also be defined for x values at which the function itself is undefined, e.g.

$$\lim_{x\to 0} \frac{\sin x}{x} = 1.$$

The function $\frac{\sin x}{x}$ is not defined for $x = 0$ but the function value approaches the limit 1 as x gets closer and closer to zero (from either side). We will learn how to calculate this limit in Chapter 6. ■

■ **Example 15.3** Consider the function $f(x) = \frac{x^2-1}{x-1}$, which is not defined at $x = 1$, but we can determine the limit

$$\lim_{x\to 1} \frac{x^2-1}{x-1} = \lim_{x\to 1} \frac{(x-1)(x+1)}{x-1} = \lim_{x\to 1}(x+1) = 2.$$

While this may seem like cheating (the numerator simply contained the denominator as a factor we could cancel), this trick will help us later. ■

15.2 The Derivative

At a certain point x of a function, one property of interest is the rate of change. For a linear function this is determined by the gradient, or slope, i.e. the ratio of the change of the function value $\Delta y = y_1 - y_0$ over the change in the independent variable $\Delta x = x_1 - x_0$,

$$\text{gradient} = \frac{y_1 - y_0}{x_1 - x_0} = \frac{f(x_0+\Delta x) - f(x_0)}{\Delta x} = \frac{(m(x_0+\Delta x)+b) - (mx_0+b)}{\Delta x} = m. \tag{15.2}$$

For such a linear function, the slope is clearly constant for any x, but we can generalize this to 'curved' functions (as long as they are sufficiently smooth) and determine an x-dependent slope, geometrically interpreted as the gradient of a tangent line at a point $(x, f(x))$ to the function. We can define the tangent gradient as the limit of secant lines between finite points that come closer and closer together. While the difference $\Delta x = x_1 - x_0$ approaches zero, the gradient defined as thus approaches a finite value if the function is well-behaved at the given point (no step or kink). This is how the derivative of a function is mathematically defined.

Definition 15.2 — Derivative of a Function. The derivative of $f(x)$ is the **slope** of, or the **gradient of the tangent** to, the function $f(x)$ at the point x. It is mathematically defined as the limit

$$\frac{df}{dx} = \lim_{\Delta x\to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}. \tag{15.3}$$

The derivative defined as such, i.e. for general x, is therefore a function in itself, $\frac{df}{dx} = \frac{df}{dx}(x)$.

R As a shorthand, we will also use the alternative notation

$$f'(x) \equiv \frac{df}{dx},$$

and similarly for higher order derivatives (see below).

To denote the gradient at a specific point x_0, we will use the following notations interchangeably:

$$f'(x_0) \equiv \frac{df}{dx}(x_0) \equiv \left.\frac{df}{dx}\right|_{x=x_0}.$$

R While it may be obvious, note that the expression $\frac{df}{dx}$ is a notation and the d's are part of it, i.e. it is not allowed to cancel them! It is generally also not allowed to break up the fraction when doing mathematical manipulations, although we will learn that this may be useful for physicists.

The actual tangent line to the function f at a specific point x_0 is then constructed as

$$y_T(x) = f(x_0) + f'(x_0)(x - x_0). \tag{15.4}$$

The **angle** θ of this tangent with the positive x-axis is therefore determined by $\tan\theta = f'(x_0)$.

It is an important application in many physics contexts that we can treat the tangent line as an **approximation** of $f(x)$ near the point x_0,

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0), \quad \text{for } x \text{ near } x_0. \tag{15.5}$$

Given that $\frac{df}{dx}$ is a function, the procedure can be repeated and we can define the second derivative of $f(x)$,

$$f''(x) = \frac{d^2 f}{dx^2} = \lim_{\Delta x \to 0} \frac{f'(x + \Delta x) - f'(x)}{\Delta x}. \tag{15.6}$$

Formula 15.3 — Higher Order Derivatives. Generally, the n-th order derivative can thus be calculated recursively as

$$f^{(n)}(x) \equiv \frac{d^n f}{dx^n} = \lim_{\Delta x \to 0} \frac{f^{(n-1)}(x + \Delta x) - f^{(n-1)}(x)}{\Delta x}, \tag{15.7}$$

where $f'(x) \equiv f^{(1)}(x)$, etc., and $f^{(0)}(x) = f(x)$ in this notation.

As noted above, the (first) derivative of a function measures the rate of change, and is graphically represented by the slope (gradient) of the tangent to the function at a given point. The 2nd derivative thus gives the rate of change of the slope, which can be interpreted in terms of the **curvature** of the original function.

We will illustrate the above definition to determine the derivative **from first principle** (i.e. using the definition) in a few examples.

■ **Example 15.4 — Derivative of x^2.** The derivative of $f(x) = x^2$ is calculated as

$$\begin{aligned}
\frac{d(x^2)}{dx} &= \lim_{\Delta x \to 0} \frac{(x + \Delta x)^2 - x^2}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{x^2 + 2x\Delta x + (\Delta x)^2 - x^2}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{\Delta x(2x + \Delta x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} (2x + \Delta x) \\
&= 2x,
\end{aligned}$$

where we applied the trick mentioned above to calculate the limit: We transformed the numerator so that we could cancel the Δx in the denominator. After this, the limit becomes easy and straightforward to determine. ■

■ **Example 15.5 — Derivative of x^n.** The above example can be generalized to calculate the derivative of $f(x) = x^n$, $(n = 1, 2, \ldots)$,

$$\frac{d(x^n)}{dx} = \lim_{\Delta x \to 0} \frac{(x + \Delta x)^n - x^n}{\Delta x}.$$

Now recall that

$$a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + a^{n-3}b^2 + \cdots + a\, b^{n-2} + b^{n-1}),$$

Applying this result to the derivative with $a = x + \Delta x$ and $b = x$ we get

$$\begin{aligned}\frac{d(x^n)}{dx} &= \lim_{\Delta x \to 0} \frac{\Delta x((x+\Delta x)^{n-1} + (x+\Delta x)^{n-2}x + (x+\Delta x)^{n-3}x^2 + \cdots + (x+\Delta x)x^{n-2} + x^{n-1})}{\Delta x} \\ &= \lim_{\Delta x \to 0} ((x+\Delta x)^{n-1} + (x+\Delta x)^{n-2}x + (x+\Delta x)^{n-3}x^2 + \cdots + (x+\Delta x)x^{n-2} + x^{n-1}) \\ &= x^{n-1} + x^{n-1} + x^{n-1} + \cdots + x^{n-1} + x^{n-1} \qquad [n \text{ terms}] \\ &= n\, x^{n-1},\end{aligned}$$

as the Δx can be simply set to zero after the fraction is cancelled. ■

R Note that this result can be extended to negative and fractional values of n - we shall assume without proof that the above result is valid for all values of n.

■ **Example 15.6 — Derivative of $\mathbf{e^x}$.** Consider first the more general function $f(x) = a^x$. The definition of the derivative gives

$$\begin{aligned}\frac{d(a^x)}{dx} &= \lim_{\Delta x \to 0} \frac{a^{x+\Delta x} - a^x}{\Delta x} \\ &= \lim_{\Delta x \to 0} a^x \frac{a^{\Delta x} - 1}{\Delta x} \\ &= a^x \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}.\end{aligned}$$

In other words, the derivative of a^x is equal to a^x times the slope of a^x at $x = 0$,

$$\frac{d(a^x)}{dx}(0) = \lim_{\Delta x \to 0} \frac{a^{0+\Delta x} - 0^x}{\Delta x} = \lim_{\Delta x \to 0} \frac{a^{\Delta x} - 1}{\Delta x}.$$

Thus we can write

$$f'(x) = a^x f'(0), \text{ for } f(x) = a^x.$$

Note that we do not get a closed result for the derivative, as it still depends on the unknown derivative at the point $x = 0$. It is natural to ask whether a value of a exists such that the derivative at $x = 0$ is $f'(0) = 1$. The irrational number for which this is the case is called **Euler's Number** $e \approx 2.71828\ldots$. Thus the function e^x is defined such that its slope at $x = 0$ equals unity and hence from the expression for the derivative of a^x we have for $f(x) = e^x$,

$$\frac{d(e^x)}{dx} = e^x,$$

i.e. e^x is a function where the slope is equal to its function value at each point x. ■

■ **Example 15.7 — Derivative of $\sin x$.** If $f = \sin x$ then from the definition of the derivative we get

$$\frac{d(\sin x)}{dx} = \lim_{\Delta x \to 0} \frac{\sin(x+\Delta x) - \sin x}{\Delta x}.$$

Using the trigonometric identity

$$\sin\alpha - \sin\beta = 2\cos\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2},$$

the limit can be written as

$$\begin{aligned}\frac{d(\sin x)}{dx} &= \lim_{\Delta\to 0}\frac{2\cos\left(x+\frac{\Delta x}{2}\right)\sin\left(\frac{\Delta x}{2}\right)}{\Delta x}\\ &= \lim_{\Delta x\to 0}\frac{\cos\left(x+\frac{\Delta x}{2}\right)\sin(\frac{\Delta x}{2})}{\frac{\Delta x}{2}}\\ &= \cos x \lim_{\Delta x\to 0}\frac{\sin(\frac{\Delta x}{2})}{\frac{\Delta x}{2}}\\ &= \cos x.\end{aligned}$$

Here we used the previously mentioned limit

$$\lim_{\frac{\Delta x}{2}\to 0}\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}} = 1.$$

A similar calculation for $f(x) = \cos x$ yields

$$\frac{d(\cos x)}{dx} = -\sin x.$$

■

16. Derivatives of Basic Functions (Riley 2.1.1)

Differentiation from first principles is important to know and understand but time-consuming in practical calculations. The general approach is to use a list of derivatives of basic functions as a basis for more complicated functions.

Formula 16.1 — Derivatives of Basic Functions. A set of useful derivatives is the following:

$$\frac{dx^n}{dx} = n\, x^{n-1}, \tag{16.1}$$

$$\frac{de^{ax}}{dx} = a\, e^{ax}, \tag{16.2}$$

$$\frac{d\ln x}{dx} = \frac{1}{x}, \tag{16.3}$$

$$\frac{d\sin(ax)}{dx} = a\cos(ax), \tag{16.4}$$

$$\frac{d\cos(ax)}{dx} = -a\sin(ax), \tag{16.5}$$

$$\frac{d\tan(ax)}{dx} = \frac{a}{\cos^2(ax)} = a(1+\tan^2(ax)), \tag{16.6}$$

$$\frac{d\arcsin(x/a)}{dx} = \frac{1}{\sqrt{a^2-x^2}}, \tag{16.7}$$

$$\frac{d\arccos(x/a)}{dx} = \frac{-1}{\sqrt{a^2-x^2}}, \tag{16.8}$$

$$\frac{d\arctan(x/a)}{dx} = \frac{a}{a^2+x^2}. \tag{16.9}$$

R While some of the above derivatives can be quite easily derived using the following rules, we recommend to learn them by heart through practice during the course. Knowing the above derivatives will speed up many of your calculations.

17. Derivatives of Compound Functions

More complicated function can in almost all cases be seen as sums, products, quotients etc. of basic functions. We here discuss the differentiation of such compound functions.

17.1 Derivative of a Sum/Difference

The simplest way to combine functions is by adding or subtracting them from each other.

Formula 17.1 — Derivative of Sum/Difference. Allowing for additional constant factors a, b, the derivative of $f(x) = au(x) \pm bv(x)$ is given by

$$\frac{df}{dx} = a\frac{du}{dx} \pm b\frac{dv}{dx}, \tag{17.1}$$

i.e. the derivative of a sum (difference) is the sum (difference) of derivatives, and constants can be factored outside the derivative. This can be easily shown using the definition of the derivative.

■ **Example 17.1** The derivative of the linear function $y(x) = mx + b$ is given by

$$\frac{d(mx+b)}{dx} = m\frac{d(x)}{dx} + b\frac{d(1)}{dx} = m, \tag{17.2}$$

as the derivative of the constant function $v(x) = 1$ is identically zero. ■

17.2 The Product Rule (Riley 2.1.2)

Consider a function $f(x)$ given as a product of two other functions,

$$f(x) = u(x)v(x). \tag{17.3}$$

The derivative of this is calculated from the definition as

$$\frac{d(u(x)v(x))}{dx} = \lim_{\Delta x \to 0} \frac{u(x+\Delta x)v(x+\Delta x) - u(x)v(x)}{\Delta x}$$
$$= \lim_{\Delta x \to 0} \frac{u(x+\Delta x)[v(x+\Delta x) - v(x)] + [u(x+\Delta x) - u(x)]v(x)}{\Delta x}, \tag{17.4}$$

where we added and subtracted $u(x+\Delta x)v(x)$, as a result of which we can write the derivative as

$$\frac{d(u(x)v(x))}{dx} = \lim_{\Delta x \to 0} \left\{ u(x+\Delta x) \left[\frac{v(x+\Delta x) - v(x)}{\Delta x} \right] + \left[\frac{u(x+\Delta x) - u(x)}{\Delta x} \right] v(x) \right\}$$
$$= u(x) \lim_{\Delta x \to 0} \left[\frac{v(x+\Delta x) - v(x)}{\Delta x} \right] + v(x) \lim_{\Delta x \to 0} \left[\frac{u(x+\Delta x) - u(x)}{\Delta x} \right]. \tag{17.5}$$

Here, we moved $u(x+\Delta x)$ and $v(x+\Delta x)$ outside the limit as they simply approach $u(x)$ and $v(x)$, respectively, as $\Delta x \to 0$. The two limits in the above equation are the definitions of the derivatives $\frac{dv}{dx}$ and $\frac{du}{dx}$, respectively.

Formula 17.2 — Product Rule of Differentiation. Thus we arrive at the product rule of differentiation,

$$\frac{d(uv)}{dx} = u\frac{dv}{dx} + v\frac{du}{dx}, \tag{17.6}$$

or more compactly using the short-hand notation,

$$(uv)' = uv' + u'v. \tag{17.7}$$

The rule can be extended to the product of three functions,

$$(uvw)' = u'vw + uv'w + uvw', \tag{17.8}$$

and similar for products of more functions.

■ **Example 17.2** Calculate the derivative of $f(x) = x^3 \sin x$.

$$\frac{d}{dx}\left(x^3 \sin x\right) = x^3 \frac{d(\sin x)}{dx} + \frac{d(x^3)}{dx} \sin x$$
$$= x^3 \cos x + 3x^2 \sin x$$
$$= x^2(x \cos x + 3 \sin x),$$

using the derivatives of basic functions. ■

■ **Example 17.3** The product rule can be used to differentiate $f(x) = x^n$ $(n = 1, 2, \ldots)$. Viewing the function as a product of n terms, $f(x) = x\,x\,\ldots\,x\,x$ yields

$$\frac{d(\overbrace{x\,x\,\ldots\,x}^{n \text{ times}})}{dx} = n \left(\overbrace{x\,x\,\ldots\,x}^{(n-1) \text{ times}} \frac{d(x)}{dx} \right)$$
$$= nx^{n-1},$$

as $(x)' = 1$ and the product rule yields a sum of the n terms. ■

17.3 The Quotient Rule (Riley 2.1.4)

In order to calculate the derivative of a quotient of functions,

$$f(x) = \frac{u(x)}{v(x)}, \tag{17.9}$$

it is easiest to differentiate the equivalent equation

$$f(x)v(x) = u(x), \tag{17.10}$$

on both sides. Using the product rule yields

$$f'(x)v(x) + f(x)v'(x) = u'(x). \tag{17.11}$$

Solving for $f'(x)$,

$$\begin{aligned} f'(x) &= \frac{u'(x) - f(x)v'(x)}{v(x)} \\ &= \frac{u'(x) - \frac{u(x)}{v(x)}v'(x)}{v(x)}, \end{aligned} \tag{17.12}$$

where we inserted $f(x) = \frac{u(x)}{v(x)}$ on the right-hand side.

Formula 17.3 — Quotient Rule of Differentiation. We thus arrive at the quotient rule of differentiation,

$$\left(\frac{u(x)}{v(x)}\right)' = \frac{u'(x)v(x) - u(x)v'(x)}{v^2(x)}. \tag{17.13}$$

Note that the derivative will not be defined for points x_0 at which $v(x_0) = 0$.

■ **Example 17.4** For $f(x) = \tan x = \frac{\sin x}{\cos x}$, the derivative is

$$\begin{aligned} \frac{d\tan x}{dx} &= \frac{\cos x \cos x - \sin x(-\sin x)}{\cos^2 x} \\ &= 1 + \frac{\sin^2 x}{\cos^2 x} \\ &= 1 + \tan^2 x = \frac{1}{\cos^2 x} = \sec^2 x, \end{aligned}$$

with $\sec x = \frac{1}{\cos x}$. ■

■ **Example 17.5** Similarly, if $f(x) = \cot x = \frac{\cos x}{\sin x}$,

$$\begin{aligned} \frac{d\cot x}{dx} &= \frac{(-\sin x)\sin x - \cos x(\cos x)}{\sin^2 x} \\ &= -\left(1 + \frac{\cos^2 x}{\sin^2 x}\right) \\ &= -(1 + \cot^2 x) = -\frac{1}{\sin^2 x} = -\csc^2 x. \end{aligned}$$

■

Example 17.6 If $f(x) = \csc x \equiv \frac{1}{\sin x}$, then

$$\begin{aligned}\frac{d\csc x}{dx} &= \frac{0-\cos x}{\sin^2 x}\\ &= \frac{\cos x}{\sin^2 x} = \frac{\cot x}{\sin x} = \cot x \csc x.\end{aligned}$$

■

Example 17.7 Similarly, if $f(x) = \sec x \equiv \frac{1}{\cos x}$, then

$$\begin{aligned}\frac{d\sec x}{dx} &= \frac{0-(-\sin x)}{\cos^2 x}\\ &= \frac{\sin x}{\cos^2 x} = \frac{\tan x}{\cos x} = \tan x \sec x.\end{aligned}$$

■

17.4 The Chain Rule (Riley 2.1.3)

Consider the situation of nested functions, i.e. where (the function value of) a function $u(x)$ is the argument of another function f, i.e. $f(u(x))$; a function inside a function. A simple example is

$$f(x) = (3+x^2)^3 = u^3(x), \tag{17.14}$$

where

$$u(x) = 3+x^2. \tag{17.15}$$

To differentiate such a function we use the definition, and multiply with $(u(x+\Delta x)-u(x))$ in the numerator and denominator,

$$\begin{aligned}\frac{df}{dx} &= \lim_{\Delta x\to 0}\frac{f(u(x+\Delta x))-f(u(x))}{\Delta x}\\ &= \lim_{\Delta x\to 0}\left\{\frac{f(u(x+\Delta x))-f(u(x))}{u(x+\Delta x)-u(x)}\frac{u(x+\Delta x)-u(x)}{\Delta x}\right\}\\ &= \lim_{\Delta u\to 0}\frac{f(u+\Delta u)-f(u)}{\Delta u}\lim_{\Delta x\to 0}\frac{u(x+\Delta x)-u(x)}{\Delta x},\end{aligned} \tag{17.16}$$

where we rewrote the argument of f as $u(x+\Delta x) = u+\Delta u$ and thus $u(x+\Delta x)-u(x) = \Delta u$ (a change of x leads to an associated change in u).

Formula 17.4 — Chain Rule of Differentiation. We can interpret the resulting limits as derivatives and thus arrive at the chain rule,

$$\frac{df(u(x))}{dx} = \frac{df}{du}\frac{du}{dx}, \tag{17.17}$$

i.e. the product of the derivative of the outer function with respect to its immediate argument (u, not x) and the derivative of the inner function with respect to x. The chain rule can be extended to more functions being nested,

$$\frac{df(u(v(x)))}{dx} = \frac{df}{du}\frac{du}{dv}\frac{dv}{dx}, \tag{17.18}$$

etc.

For the above example, the chain rule yields

$$\begin{aligned}
\frac{d}{dx}\left((3+x^2)^3\right) &= \frac{d(u^3)}{du}\frac{d(3+x^2)}{dx} \\
&= 3u^2\,2x \\
&= 3\left(3+x^2\right)^2\,2x = 6x\left(3+x^2\right)^2. \qquad (17.19)
\end{aligned}$$

Note that on the last line we inserted $u = 3+x^2$ to make the final result an explicit function of x only.

R There is a subtle issue when writing a nested function suggestively as $f(x) = f(u(x))$: Strictly speaking the function name on the left-hand side of this identity should be different from f, e.g. $\bar{f}(x) = f(u(x))$. This is because the nesting of u into f results in a different functional dependence from that of f. For example consider

$$\begin{aligned}
&u(x) = x^2, \text{ and } f(u) = 1+u, \\
\Rightarrow\ &\bar{f}(x) = f(u(x)) = 1+x^2.
\end{aligned}$$

Clearly, f is a linear function whereas $\bar{f}(x)$ is a different, quadratic function, and for consistency one should not use the same function name to identify them.

■ **Example 17.8** Consider $\bar{f}(x) = e^{ax}$ (one of the basic functions),

$$\begin{aligned}
&f(u) = e^u, \quad \text{and} \quad u(x) = ax, \\
&\frac{df}{du} = e^u, \quad \text{and} \quad \frac{du}{dx} = a, \\
\Rightarrow\ &\frac{d(e^{ax})}{dx} = ae^u = ae^{ax}.
\end{aligned}$$

■

■ **Example 17.9** Consider $\bar{f}(x) = \sin^3 x \equiv (\sin x)^3$,

$$\begin{aligned}
&f(u) = u^3, \quad \text{and} \quad u(x) = \sin x, \\
&\frac{df}{du} = 3u^2, \quad \text{and} \quad \frac{du}{dx} = \cos x, \\
\Rightarrow\ &\frac{d\bar{f}}{dx} = 3u^2\cos x = 3\sin^2 x\,\cos x.
\end{aligned}$$

■

■ **Example 17.10** On the other hand, consider $\bar{f}(x) = \sin(x^3)$ where the same functions are nested in different order,

$$\begin{aligned}
&f(u) = \sin u, \quad \text{and} \quad u(x) = x^3, \\
&\frac{df}{du} = \cos u, \quad \text{and} \quad \frac{du}{dx} = 3x^2, \\
\Rightarrow\ &\frac{d\bar{f}}{dx} = \cos u\; 3x^2 = 3x^2\cos(x^3).
\end{aligned}$$

■

■ **Example 17.11** Consider $\bar{f}(x) = \frac{1}{u(x)}$,

$$f(u) = \frac{1}{u}, \quad \text{and} \quad u(x),$$
$$\frac{df}{du} = \frac{-1}{u^2}, \quad \text{and} \quad \frac{du}{dx},$$
$$\Rightarrow \frac{d\bar{f}}{dx} = -\frac{1}{u^2}\frac{du}{dx} = -\frac{u'(x)}{u^2(x)}.$$

This coincides with the result one would get using the quotient rule. ■

■ **Example 17.12** Consider $\bar{f}(x) = \ln(u(x))$,

$$f(u) = \ln u, \quad \text{and} \quad u(x),$$
$$\frac{df}{du} = \frac{1}{u}, \quad \text{and} \quad \frac{du}{dx},$$
$$\Rightarrow \frac{d\bar{f}}{dx} = \frac{1}{u}\frac{du}{dx} = \frac{u'(x)}{u(x)}.$$

■

17.5 Inverse Differentiation

Definition 17.5 If $y = f(x)$ describes a function, the **inverse function** is written as $x = f^{-1}(y)$. It is defined by

$$f^{-1}(f(x)) = f(f^{-1}(x)) = x, \tag{17.20}$$

nesting one into the other results in the original argument, x.

Formula 17.6 — Rule of Inverse Differentiation. Differentiating this as per the chain rule thus leads to the relation

$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}}. \tag{17.21}$$

This rule can be very useful when differentiating a function for which the derivative of the inverse function is well known. We illustrate this for a couple of examples.

■ **Example 17.13** To differentiate the **natural logarithmic function** $y = \ln x$, we note that its inverse function is the natural exponential, $x = e^y$, because $e^{\ln x} = x$. The derivative of $\ln x$ is thus

$$\frac{d\ln x}{dx} = \frac{dy}{dx} = \frac{1}{\frac{dx}{dy}} = \frac{1}{e^y}.$$

We now have to express the right-hand side in terms of x by substituting $y = \ln x$,

$$\frac{d\ln x}{dx} = \frac{1}{x}$$

■

■ **Example 17.14** Differentiate the inverse sine function $y = \sin^{-1} x \equiv \arcsin x$.

Clearly, $x = \sin y$ and thus

$$\frac{dy}{dx} = \frac{1}{\frac{dx}{dy}} = \frac{1}{\cos y}.$$

To express the right-hand side in terms of x, we first use $\sin^2 y + \cos^2 y = 1$, i.e. $\cos y = \sqrt{1-\sin^2 y}$ and therefore

$$\begin{aligned}\frac{d\arcsin x}{dx} &= \frac{1}{\sqrt{1-\sin^2 y}}\\ &= \frac{1}{\sqrt{1-x^2}}.\end{aligned}$$

A similar calculation for $y = \cos^{-1} x \equiv \arccos x$ yields

$$\frac{d\arccos x}{dx} = \frac{-1}{\sqrt{1-x^2}}.$$

■

The hyperbolic sine and cosine functions were already defined before,

$$\sinh x = \frac{1}{2}(e^x - e^{-x}), \tag{17.22}$$

$$\cosh x = \frac{1}{2}(e^x + e^{-x}). \tag{17.23}$$

From this, we readily find the relations

$$\cosh x + \sinh x = e^x, \tag{17.24}$$

$$\cosh^2 x - \sinh^2 x = 1. \qquad \text{[Note the minus sign!]} \tag{17.25}$$

Furthermore, the hyperbolic tangent function is defined as expected,

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \tag{17.26}$$

The derivatives of the hyperbolic functions are easily calculated,

$$\frac{d\sinh x}{dx} = \frac{1}{2}(e^x + e^{-x}) = \cosh x. \tag{17.27}$$

$$\frac{d\cosh x}{dx} = \frac{1}{2}(e^x - e^{-x}) = \sinh x. \tag{17.28}$$

$$\frac{d\tanh x}{dx} = 1 - \tanh^2 x = \frac{1}{\cosh^2 x}. \tag{17.29}$$

■ **Example 17.15** Among the inverse hyperbolic functions we take $y = \sinh^{-1}\left(\frac{x}{a}\right)$ as an example. Then $x = a\sinh y$ and

$$\frac{d}{dx}\left(\sinh^{-1}\left(\frac{x}{a}\right)\right) = \frac{dy}{dx} = \frac{1}{\frac{dx}{dy}} = \frac{1}{a\cosh y} = \frac{1}{a\sqrt{\sinh^2 y + 1}} = \frac{1}{\sqrt{x^2+a^2}}.$$

■

18. Stationary Points and Extrema (Riley 2.1.8)

Derivatives give the rate of change of a function. An important application of this is finding **local maxima and minima** of functions.

Formula 18.1 — Stationary Point of a Function. If at some point x_0 we have

$$f'(x_0) = 0, \tag{18.1}$$

then this is a **stationary point**, the tangent at this point is horizontal and the function value is approximately constant for small deviations from x_0.

For a local maximum or minimum, $f'(x)$ needs to change sign around x_0, i.e. the **curvature** $f''(x_0)$ at this point has to be negative or positive, respectively. The nature of a stationary point x_0 ($f'(x_0) = 0$) can be:

1. **Local Minimum** if $f''(x_0) > 0$ (positive curvature).
2. **Local Maximum** if $f''(x_0) < 0$ (negative curvature).
3. **Point of Inflection** if $f''(x_0) = 0$ and f'' changes sign through the point.
4. **Undetermined** in any other case, i.e. specifically if $f''(x_0) = 0$ and f'' does not change sign at x_0. It will be necessary to consider even higher derivatives or inspect the function.

R We emphasize the qualification that the **extrema** (maxima and minima) determined as above are of a local nature only, i.e. they are the points with the largest or smallest function value compared to their neighbourhood. To determine whether they are a global maximum or minimum (the largest or smallest function value of any possible x) requires further analysis.

■ **Example 18.1** Find the stationary point(s) of

$$f(x) = x \ln x, \quad x > 0.$$

and determine its (their) nature.

For the derivative we have

$$f'(x) = \ln x + x\frac{1}{x} = \ln x + 1.$$

To find the stationary point(s) we solve $f'(x) = 0$, i.e. $\ln x = -1$. Thus there is only one stationary point at

$$x_0 = e^{-1} = \frac{1}{e} \approx 0.37.$$

To determine its nature we have to calculate the second derivative,

$$f''(x) = \frac{1}{x},$$

and evaluate it at $x_0 = 1/e$,

$$f''(x_0) = e > 0.$$

Thus the stationary point is a minimum. It is in fact the global minimum of the function. ■

■ **Example 18.2** Find the stationary point(s) of $f(x) = x^4$ and determine its (their) nature.
We start by calculating the derivative:

$$f'(x) = 4x^3,$$

thus $x_0 = 0$ is the only stationary point. To determine its nature, we calculate the 2nd derivative,

$$f''(x) = 12x^2,$$

and thus $f''(0) = 0$. However, $f''(x)$ does not change sign around $x = 0$ and the nature of the stationary point remains undetermined according to our criteria. In this instance, it is rather easy to see by inspection that $x_0 = 0$ is a (local and global) minimum of $f(x) = x^4$. ■

19. Implicit and Parametric Differentiation

19.1 Implicit Differentiation (Riley 2.1.5)

So far we have only considered functions where the dependent variable (the function value) is given explicitly in terms of the independent variable x, i.e. $y = f(x)$. In other words, only the variable x appears on the right-hand side. This may not be always possible (one cannot solve for $y(x)$) or desirable. Instead, we may have a situation where $y(x)$ and x are **implicitly** connected.

Definition 19.1 — Implicit Relation. In general we can write this as

$$f(x, y(x)) = \text{const.}, \tag{19.1}$$

defining an implicit relation between x and $y(x)$. Here const. simply represents a constant value and any dependence on x and y is contained on the left-hand side in f by convention.

The calculation of the derivative $\frac{dy}{dx}$ in such a case is best illustrated in an example:

$$f(x, y(x)) = x^3 - 3xy(x) + y^3(x) = 2. \tag{19.2}$$

To proceed, we differentiate this expression term by term with respect to x,

$$\frac{d}{dx}\left(x^3 - 3xy(x) + y^3(x)\right) = \frac{d}{dx}(2), \tag{19.3}$$

$$\Rightarrow \frac{d}{dx}(x^3) - \frac{d}{dx}(3xy(x)) + \frac{d}{dx}(y^3(x)) = 0, \tag{19.4}$$

$$\Rightarrow 3x^2 - \left(3y(x) + 3x\frac{dy}{dx}\right) + 3y^2(x)\frac{dy}{dx} = 0. \tag{19.5}$$

This is called **implicit differentiation**. We then solve the resulting equation for $\frac{dy}{dx}$,

$$\frac{dy}{dx} = \frac{y(x) - x^2}{y^2(x) - x}. \tag{19.6}$$

Unless we insert an explicit form of y in terms of x, the right-hand side of this equation will depend on both x and y.

■ **Example 19.1** A circle with unit radius around the origin is described by

$$x^2 + y^2 = 1.$$

To find the slope of the tangent at a given point (x, y) on the circle, we can use implicit differentiation,

$$\frac{d}{dx}(x^2) + \frac{d}{dx}(y^2) = 2x + 2y\frac{dy}{dx} = 0,$$

and thus

$$\frac{dy}{dx} = -\frac{x}{y}.$$

For example, the point $x = 1/2$, $y = \sqrt{3}/2$ is situated on this circle, and thus the slope of the tangent to the circle at this point is $\frac{dy}{dx} = -1/\sqrt{3}$. The tangent thus has an angle of $\theta = \arctan(-1/\sqrt{3}) = -\pi/6$ against the horizontal axis. ■

19.2 Parametric Differentiation

Yet another way to prescribe a relation between coordinates x and y is through their dependence on a shared parameter (e.g. time in a physical example),

$$x = x(t), \tag{19.7}$$

$$y = y(t). \tag{19.8}$$

This describes a **parametrized curve** in 2D space. The parametric equation of a line is an example for this (in 3D).

Formula 19.2 — Parametric Derivative. To calculate the derivative $\frac{dy}{dx}$ in this situation, we make use of the chain rule and the derivative of an inverse function,

$$\frac{dy(t(x))}{dx} = \frac{dy}{dt}\frac{dt}{dx} = \frac{dy}{dt}\left(\frac{dx}{dt}\right)^{-1}. \tag{19.9}$$

■ **Example 19.2** Consider the curve defined by:

$$x(t) = 2\cos t, \quad y(t) = 2\sin t, \quad 0 \leq t < 2\pi.$$

It is easy to see that this describes a circle. Indeed, we can eliminate the parameter t by dividing both parametric equations by 2, and by adding the squares,

$$\left(\frac{x}{2}\right)^2 + \left(\frac{y}{2}\right)^2 = \cos^2 t + \sin^2 t = 1.$$

Thus x and y satisfy the the equation of a circle with radius = 2,

$$x^2 + y^2 = 2^2.$$

■

■ **Example 19.3** The parametric equations

$$x(t) = 2t, \quad y(t) = 4t^2, \quad -\infty < t < +\infty,$$

describe a parabola. Indeed, eliminating t we obtain

$$y = 4\left(\frac{x}{2}\right)^2 = x^2,$$

which is the standard equation of a parabola. ■

■ **Example 19.4** The coordinates of a moving vehicle are given by

$$x(t) = -t^2, \quad y(t) = (1/3)t^3,$$

where t is time. Find dy/dx when $t = 2$.

We thus need to calculate the parametric derivative, and evaluate it at $t = 2$,

$$\left.\frac{dy}{dx}\right|_{t=2}.$$

The general parametric derivative is

$$\frac{dy}{dx} = \frac{dy}{dt}\frac{dt}{dx} = \frac{dy}{dt}\left(\frac{dx}{dt}\right)^{-1} = t^2(-2t)^{-1} = -\frac{t}{2},$$

which for $t = 2$ is

$$\left.\frac{dy}{dx}\right|_{t=2} = -1.$$

The vehicle travels in the exact north-west direction (identifying positive x direction = east and positive y direction = north) at $t = 2$. ■

■ **Example 19.5** While the above examples illustrate the principle, parametric differentiation is especially useful when it is not possible to solve for the parameter, i.e. express the curve in the explicit form $y(x)$. Consider the parametric curve

$$x(t) = t\cos(\pi t), \quad y(t) = t\sin(\pi t), \quad 0 \le t < +\infty.$$

This describes a spiral starting at $x = 0$, $y = 0$, with an increasing radius $r(t) = t$ (e.g. t is time and the curve describes the motion of a particle). The spiral makes a full cycle after $t = 2, 4, 6, \ldots$. It is not possible to write down explicitly $y(x)$ for all of the points of the curve and to calculate $\frac{dy}{dx}$ we use parametric differentiation,

$$\frac{dy}{dx} = \frac{dy}{dt}\left(\frac{dx}{dt}\right)^{-1} = \frac{\sin(\pi t) + \pi t\cos(\pi t)}{\cos(\pi t) - \pi t\sin(\pi t)}.$$

For example, at $t = 0$, the slope is

$$\left.\frac{dy}{dx}\right|_{t=0} = 0,$$

i.e. a particle following such a path in time t travels eastbound at the very start of the spiral. This is very different from the corresponding motion on a circle with a constant radius where the initial movement is northbound. On the other hand, when t is large, the 2nd terms in the numerator and denominator are dominant, and thus

$$\left.\frac{dy}{dx}\right|_{t\text{ large}} \to \frac{\pi t\cos(\pi t)}{-\pi t\sin(\pi t)} = -\frac{\cos(\pi t)}{\sin(\pi t)} = -\frac{x}{y}.$$

The is the same as for a circle, compare with Example 19.1. ■

20. Vector Differentiation (Riley 10.1)

We can extend the notion of differentiation to vector-valued functions, i.e. a vector depends on a single variable t, $\mathbf{a} = \mathbf{a}(t)$. The parametrization of a line can be seen as an example of this: $\mathbf{r}(\lambda) = \mathbf{r}_0 + \lambda\mathbf{b}$ is a linear function of λ with constant vectors $\mathbf{r}_0$ and $\mathbf{b}$.

Definition 20.1 — Derivative of a Vector-valued Function. Given a vector-valued function $\mathbf{a} = \mathbf{a}(t)$, its derivative with respect to t is defined as

$$\frac{d\mathbf{a}(t)}{dt} = \lim_{\Delta t \to 0} \frac{\mathbf{a}(t + \Delta t) - \mathbf{a}(t)}{\Delta t}. \tag{20.1}$$

The derivative $\frac{d\mathbf{a}}{dt}$ is itself a vector, in general again a function of t.

In Cartesian coordinates the function $\mathbf{a} = \mathbf{a}(t)$ can be written as

$$\mathbf{a}(t) = a_x(t)\mathbf{i} + a_y(t)\mathbf{j} + a_z(t)\mathbf{k}, \tag{20.2}$$

i.e. the components $a_x(t)$, $a_y(t)$, $a_z(t)$ are functions of t, whereas the Cartesian unit vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are constant. The derivative can then be expressed as

$$\frac{d\mathbf{a}}{dt} = \frac{da_x}{dt}\mathbf{i} + \frac{da_y}{dt}\mathbf{j} + \frac{da_z}{dt}\mathbf{k}. \tag{20.3}$$

Formula 20.2 — Properties of Vector Derivatives. Given two vector-valued functions $\mathbf{a} = \mathbf{a}(t)$

and $\mathbf{b} = \mathbf{b}(t)$, the following rules apply:

$$\frac{d(c\mathbf{a})}{dt} = c\frac{d\mathbf{a}}{dt}, \qquad [c \text{ is a constant number}], \tag{20.4}$$

$$\frac{d(f\mathbf{a})}{dt} = \frac{df}{dt}\mathbf{a} + f\frac{d\mathbf{a}}{dt}, \qquad [f \text{ is a function } f = f(t)], \tag{20.5}$$

$$\frac{d(\mathbf{a}+\mathbf{b})}{dt} = \frac{d\mathbf{a}}{dt} + \frac{d\mathbf{b}}{dt}, \tag{20.6}$$

$$\frac{d(\mathbf{a}\cdot\mathbf{b})}{dt} = \mathbf{a}\cdot\frac{d\mathbf{b}}{dt} + \frac{d\mathbf{a}}{dt}\cdot\mathbf{b}, \tag{20.7}$$

$$\frac{d(\mathbf{a}\times\mathbf{b})}{dt} = \mathbf{a}\times\frac{d\mathbf{b}}{dt} + \frac{d\mathbf{a}}{dt}\times\mathbf{b}. \tag{20.8}$$

These rules can be derived using the definition of the vector differentiation and the properties of vector products etc. Alternatively, they can also be derived by expressing the vector expressions in terms of Cartesian components of the vectors.

For example, if

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k} = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} \tag{20.9}$$

is the position vector of a particle as a function of time t, then the velocity and acceleration vectors are given by

$$\mathbf{v}(t) = \frac{d\mathbf{r}(t)}{dt} = \frac{dx}{dt}\mathbf{i} + \frac{dy}{dt}\mathbf{j} + \frac{dz}{dt}\mathbf{k} = \begin{pmatrix} dx/dt \\ dy/dt \\ dz/dt \end{pmatrix}, \tag{20.10}$$

and

$$\mathbf{a}(t) = \frac{d\mathbf{v}(t)}{dt} = \frac{d^2\mathbf{r}(t)}{dt^2} = \frac{d^2x}{dt^2}\mathbf{i} + \frac{d^2y}{dt^2}\mathbf{j} + \frac{d^2z}{dt^2}\mathbf{k} = \begin{pmatrix} d^2x/dt^2 \\ d^2y/dt^2 \\ d^2z/dt^2 \end{pmatrix}, \tag{20.11}$$

respectively.

■ **Example 20.1** A particle experiencing no force travels with constant velocity $\mathbf{v}$ in a straight line, i.e. its position vector as a function of time t is given by

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

where $\mathbf{r}_0$ is the starting position at $t = 0$. Differentiating this shows that the vector $\mathbf{v}$ is indeed the velocity vector,

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \frac{d}{dt}(\mathbf{r}_0 + t\mathbf{v}),$$

where $\mathbf{v}$ is a constant vector (does not depend on t) and thus the acceleration vector is zero, $\mathbf{a} = \frac{d\mathbf{v}}{dt} = \mathbf{0}$. ■

■ **Example 20.2** For

$$\mathbf{w}(t) = 3t^2\mathbf{i} + \cos(2t)\mathbf{j},$$

find

$$\frac{d\mathbf{w}}{dt}, \quad \left|\frac{d\mathbf{w}}{dt}\right|, \quad \frac{d^2\mathbf{w}}{dt^2}.$$

With the Cartesian unit vectors being constant, the vector derivative is easily calculated as

$$\frac{d\mathbf{w}}{dt} = 6t\mathbf{i} - 2\sin(2t)\mathbf{j} = \begin{pmatrix} 6t \\ -2\sin(2t) \\ 0 \end{pmatrix}.$$

Its magnitude is

$$\left|\frac{d\mathbf{w}}{dt}\right| = \sqrt{(6t)^2 + (-2\sin(2t))^2} = 2\sqrt{9t^2 + \sin^2(2t)},$$

and the 2nd derivative is given by

$$\frac{d^2\mathbf{w}}{dt^2} = 6\mathbf{i} - 4\cos(2t)\mathbf{j} = \begin{pmatrix} 6 \\ -4\cos(2t) \\ 0 \end{pmatrix}.$$

■

■ **Example 20.3** For a general vector function $\mathbf{r} = \mathbf{r}(t)$, show that

$$\frac{d(\mathbf{r} \times \mathbf{r}')}{dt} = \mathbf{r} \times \mathbf{r}'',$$

where $\mathbf{r}' = \mathbf{r}'(t)$ and $\mathbf{r}'' = \mathbf{r}''(t)$ are the first and second derivative of $\mathbf{r}$, respectively.

Using Equation (20.8), we have

$$\begin{aligned}\left(\mathbf{r} \times \mathbf{r}'\right)' &= \mathbf{r}' \times \mathbf{r}' + \mathbf{r} \times \mathbf{r}'' \\ &= \mathbf{r} \times \mathbf{r}'',\end{aligned}$$

because $\mathbf{u} \times \mathbf{u} = \mathbf{0}$ for any vector $\mathbf{u}$, thus $\mathbf{r}'(t) \times \mathbf{r}'(t) = \mathbf{0}$ for any value of t. ■

■ **Example 20.4** Consider the motion of a particle in a circle around the origin with radius R at a constant speed (magnitude of velocity vector) V. Show that i) the velocity vector $\mathbf{v}(t)$ is perpendicular to the position vector $\mathbf{r}(t)$ of the particle; ii) the acceleration vector $\mathbf{a}(t)$ is perpendicular to $\mathbf{v}(t)$; and iii) the magnitude of the acceleration is

$$|\mathbf{a}| = \frac{|\mathbf{v}|^2}{|\mathbf{r}|} = \frac{V^2}{R}.$$

Without specifying the exact time-dependence of the motion, we notice that while $\mathbf{r}(t)$ and $\mathbf{v}(t)$ are not constant, their magnitudes are:

$$\begin{aligned}|\mathbf{r}(t)|^2 &= \mathbf{r}(t) \cdot \mathbf{r}(t) = R^2 = \text{constant in time}, \\ |\mathbf{v}(t)|^2 &= \mathbf{v}(t) \cdot \mathbf{v}(t) = V^2 = \text{constant in time}.\end{aligned}$$

This implies:

$$\begin{aligned}\frac{d}{dt}(\mathbf{r} \cdot \mathbf{r}) = 0 \quad &\Rightarrow \quad \frac{d\mathbf{r}}{dt} \cdot \mathbf{r} + \mathbf{r} \cdot \frac{d\mathbf{r}}{dt} = 0 \\ &\Rightarrow \quad \mathbf{r} \cdot \mathbf{v} = 0 \\ &\Rightarrow \quad \mathbf{r},\ \mathbf{v} \text{ are perpendicular for any } t.\end{aligned}$$

In the same way:

$$\begin{aligned} \frac{d}{dt}(\mathbf{v}\cdot\mathbf{v}) = 0 \quad &\Rightarrow \quad \frac{d\mathbf{v}}{dt}\cdot\mathbf{v}+\mathbf{v}\cdot\frac{d\mathbf{v}}{dt} = 0 \\ &\Rightarrow \quad \mathbf{v}\cdot\mathbf{a} = 0 \\ &\Rightarrow \quad \mathbf{v},\ \mathbf{a} \text{ are perpendicular for any } t. \end{aligned}$$

Combining both results above, we can also infer that **r** and **a** are either parallel or anti-parallel. Finally we calculate

$$\frac{d}{dt}(\mathbf{r}\cdot\mathbf{v}) = \frac{d\mathbf{r}}{dt}\cdot\mathbf{v}+\mathbf{r}\cdot\frac{d\mathbf{v}}{dt} = \mathbf{v}\cdot\mathbf{v}+\mathbf{r}\cdot\mathbf{a} = 0.$$

This implies that $\mathbf{r}\cdot\mathbf{a} = -V^2$ and therefore **r** and **a** are anti-parallel, i.e. $\mathbf{r}\cdot\mathbf{a} = -|\mathbf{r}||\mathbf{a}| = -R|\mathbf{a}|$. Therefore

$$0 = \mathbf{v}\cdot\mathbf{v}+\mathbf{r}\cdot\mathbf{a} = V^2 - R|\mathbf{a}| \quad \Rightarrow \quad |\mathbf{a}| = \frac{V^2}{R} = \text{constant in time.}$$

■

The concept of the velocity vector in physics can be geometrically generalized as a tangential vector to a curve in 3D space.

Definition 20.3 In general, for any parametrized curve, i.e. described by the position vector $\mathbf{r} = \mathbf{r}(\lambda)$ as a function of a parameter λ, the vector

$$\mathbf{T}(\lambda) = \frac{d\mathbf{r}}{d\lambda} \tag{20.12}$$

is called the **tangent vector** to the curve and the normalized vector

$$\hat{\mathbf{T}}(\lambda) = \frac{\mathbf{T}(\lambda)}{|\mathbf{T}(\lambda)|}$$

is called the **unit tangent vector**. Both the tangent and unit tangent vector will generally depend on λ. As suggested by their name, they can be geometrically interpreted as vectors tangential (parallel) to the curve at a given point $\mathbf{r}(\lambda)$.

■ **Example 20.5** Find the unit tangent vector to the helix described by $x(\theta) = \cos\theta$, $y(\theta) = \sin\theta$, $z(\theta) = \theta$ with $-\infty < \theta < \infty$. Determine the equation of the tangent to the helix at the point where it crosses the xy-plane.

The helix is defined as the curve

$$\mathbf{r}(\theta) = \cos\theta\,\mathbf{i}+\sin\theta\,\mathbf{j}+\theta\,\mathbf{k} = \begin{pmatrix}\cos\theta\\ \sin\theta\\ \theta\end{pmatrix},$$

as a function of θ. The tangent and unit tangent vectors are

$$\mathbf{T}(\theta) = \frac{d\mathbf{r}}{d\theta} = \begin{pmatrix}-\sin\theta\\ \cos\theta\\ 1\end{pmatrix}, \qquad \hat{\mathbf{T}}(\theta) = \frac{1}{\sqrt{2}}\begin{pmatrix}-\sin\theta\\ \cos\theta\\ 1\end{pmatrix}.$$

The crossing of the helix with the xy-plane is defined by $z = 0$, i.e. $\theta = 0$. The location of the crossing point is therefore

$$\mathbf{r}(\theta = 0) = \begin{pmatrix}\cos 0\\ \sin 0\\ 0\end{pmatrix} = \begin{pmatrix}1\\ 0\\ 0\end{pmatrix},$$

and the unit tangent vector at this point is

$$\hat{\mathbf{T}}(\theta=0)=\frac{1}{\sqrt{2}}\begin{pmatrix}-\sin 0\\ \cos 0\\ 1\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}0\\ 1\\ 1\end{pmatrix}.$$

The equation of the tangent line, parametrized by the real parameter λ is thus given by

$$\mathbf{r}_T=\mathbf{r}(\theta=0)+\lambda\hat{\mathbf{T}}(\theta=0)=\begin{pmatrix}1\\ 0\\ 0\end{pmatrix}+\frac{\lambda}{\sqrt{2}}\begin{pmatrix}0\\ 1\\ 1\end{pmatrix}.$$

■

Exercise 20.1 Consider the set of all possible tangent lines to the above helix, i.e. a tangent line attached to each point $\mathbf{r}(\theta)$ of the helix,

$$\mathbf{r}_{T_\theta}=\mathbf{r}(\theta)+\lambda\hat{\mathbf{T}}(\theta),\quad -\infty<\theta<\infty.$$

Do any two of these tangent lines intersect with each other? ■