DIFFERENTIATION, OPTIMISATION AND INTEGRATION

INTRODUCTION TO MATHEMATICS IN DATA SCIENCE

1. DIFFERENTIATION

The main purpose of the derivative is to understand *the (instantaneous) rate of change* of a given function with respect to its variable. How would one define and compute such a thing?
A good example to have in mind when thinking of the derivative is *velocity*. Simply put, the velocity is defined as

$$\text{velocity} = \frac{\text{distance traversed}}{\text{time it took us to traverse that distance}}.$$

For example: If it took us 2 hours to traverse 100ml we would say that we have driven *with an average velocity* of 50mph. That, however, doesn't tell us the velocity at each given point of our journey, which might be very important if we want to make sure that we didn't break any speed limits. How then, could we compute our velocity at a certain given time?
Say that $x(t)$ is our position at time t, and we're interested in computing the velocity[1] at a given time t_0. Assuming that we are at times that are *very close to* t_0 we "expect" that our velocity is "almost constant" and equals the velocity at time t_0, $v(t_0)$. This means that if t is very close to t_0

$$x(t) \approx \underbrace{x(t_0)}_{\text{position at time } t_0} + \underbrace{v(t_0)(t-t_0)}_{\text{additional distance traversed in time } t-t_0}$$

or in other words

$$\frac{x(t)-x(t_0)}{t-t_0} \approx v(t_0).$$

The above approximation becomes more and more accurate as t gets closer and closer to t_0 *but does not touch it.* Thus, if we want to be able to define the derivative precisely we must *take t to t_0*, or in other words: Use the notion of a limit for the appropriate "average rate".
This is exactly the idea of how to define the rate of change of a given function (position or not).

[1] The velocity can be negative, indicating that we go backwards. The speed, which is the absolute value of the velocity, is always non-negative.

Definition 1. Let $f : (a, b) \to \mathbb{R}$ be a given function, and let x_0 be a point in (a, b). We say that f is differentiable at the point x_0 if

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

exists and is finite. If that happens, we denote the limit by $f'(x_0)$ and call it *the derivative of* f at the point x_0.

Remark 1. Looking at the definition, and how we conceived it, we can actually give a geometric interpretation to the derivative. Much like with velocity, a function that is differentiable at x_0 satisfies

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0).$$

The above means that we can approximate the function f at the point x_0 with the line

$$y = f(x_0) + f'(x_0)(x - x_0)$$

which passes through the point $\left(x_0, f(x_0)\right)$. such a line is called *the tangent line* of f at the point x_0, and the derivative, $f'(x_0)$, is exactly its slope!

Now that we have defined the derivative, we ask ourselves if we can find techniques that will help us compute it without using the definition. The next theorem, and the table that follows it, will be our main tools to compute derivatives in this module.

Theorem 1.1 (Rules of Differentiation)**.** *Let f and g be real valued functions. Then*

(i) (Linearity) If f and g are differentiable at x_0, then for any $\alpha, \beta \in \mathbb{R}$ the function $\alpha f(x) + \beta g(x)$ is also differentiable at x_0 and

$$\left(\alpha f + \beta g\right)'(x_0) = \alpha f'(x_0) + \beta g'(x_0).$$

(ii) (Product Rule) If f and g are differentiable at x_0, then the function $h(x) = f(x) \cdot g(x)$ is also differentiable at x_0 and

$$h'(x_0) = \left(f \cdot g\right)'(x_0) = f'(x_0)g(x_0) + g'(x_0)f(x_0).$$

(iii) (Quotient Rule) If f and g are differentiable at x_0 and $g(x_0) \neq 0$, then the function $h(x) = \frac{f(x)}{g(x)}$ is also differentiable at x_0 and

$$h'(x_0) = \left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - g'(x_0)f(x_0)}{\left(g(x_0)\right)^2}.$$

(iv) (Chain Rule) If g is differentiable at x_0 and f is differentiable at $g(x_0)$, then the function $h(x) = \left(f \circ g\right)(x) = f\left(g(x)\right)$ is also differentiable at x_0 and

$$h'(x_0) = \left(f \circ g\right)'(x_0) = f'\left(g(x_0)\right)g'(x_0).$$

Table of Derivatives

Function	Derivative
$f(x) = C,\ C \in \mathbb{R}$	$f'(x) = 0$
$f(x) = x$	$f'(x) = 1$
$f(x) = x^n,\ n \in \mathbb{N}$	$f'(x) = nx^{n-1}$
$f(x) = x^\alpha,\ \alpha \in \mathbb{R}$	$\alpha x^{\alpha-1}$
$f(x) = \sin(x)$	$f'(x) = \cos(x)$
$f(x) = \cos(x)$	$f'(x) = -\sin(x)$
$f(x) = \tan(x)$	$f'(x) = \frac{1}{\cos^2(x)}$
$f(x) = \cot(x)$	$f'(x) = -\frac{1}{\sin^2(x)}$
$f(x) = e^x$	$f'(x) = e^x$
$f(x) = a^x,\ a > 0$	$f'(x) = a^x \ln(a)$
$f(x) = \ln(x)$	$f'(x) = \frac{1}{x}$

Example 1. Find the derivatives of the following functions:

a) $f(x) = x^2 + 3x + 2$.
b) $f(x) = \frac{x}{x-1}$.
c) $f(x) = x^4 + \sin(x)$.
d) $f(x) = \frac{1}{2+\cos(x)}$.
e) $f(x) = x^4 \sin(x)$.
f) $f(x) = \frac{2-\sin(x)}{2+\sin(x)}$.

Solution. a) Using the linearity of the derivative we find that

$$f'(x) = 2x + 3 + 0 = 2x + 3.$$

b) Using the quotient rule we find that

$$f'(x) = \frac{1 \cdot (x-1) - x \cdot 1}{(x-1)^2} = -\frac{1}{(x-1)^2}.$$

c) Using the linearity of the derivative we find that

$$f'(x) = 4x^3 + \cos(x) + 3 + 0 = 2x + 3.$$

d) Using the quotient rule we find that

$$f'(x) = \frac{0 \cdot (2 + \cos(x)) - 1 \cdot (0 - \sin(x))}{(2 + \cos(x))^2} = \frac{\sin(x)}{(2 + \cos(x))^2}.$$

e) using the product rule we find that

$$f'(x) = 4x^3 \cdot \sin(x) + x^4 \cdot \cos(x).$$

f) Using the quotient rule we find that

$$f'(x) = \frac{(2 - \cos(x))(2 + \sin(x)) - (2 - \sin(x))(2 + \cos(x))}{(2 + \sin(x))^2}$$

$$= \frac{4-2\cos(x)+2\sin(x)-\sin(x)\cos(x)-(4-2\sin(x)+2\cos(x)-\sin(x)\cos(x))}{(2+\sin(x))^2}$$

$$= \frac{4(\sin(x)-\cos(x))}{(2+\sin(x))^2}$$

□

Example 2. Find the derivatives of the following functions:

a) $f(x) = \cos(2x) - \sin x$.
b) $f(x) = \sqrt{1+x^2}$.
c) $f(x) = \left(2-x^2\right)\cos\left(x^2\right)$.
d) $f(x) = \sin(\sin(x))$.
e) $f(x) = \frac{x}{\sqrt{4-x^2}}$.

Solution. a) We start by differentiating $\cos(2x)$. Denoting by $h(x) = \cos(x)$ and $g(x) = 2x$ we see that

$$\cos(2x) = h\left(g(x)\right)$$

and as such, using the chain rule, we find that

$$(\cos(2x))' = h'(g(x))g'(x) = -\sin(2x)\cdot 2 = -2\sin(2x).$$

Thus, with the linearity of the derivative we find that

$$f'(x) = -2\sin(2x) - \cos(x).$$

b) Denoting by $h(x) = \sqrt{x} = x^{\frac{1}{2}}$ and $g(x) = 1 + x^2$ we see that

$$f(x) = h\left(g(x)\right)$$

and as such, using the chain rule, we find that

$$f'(x) = h'(g(x))g'(x) = \frac{1}{2}\left(1+x^2\right)^{-\frac{1}{2}}\cdot 2x = \frac{x}{\sqrt{1+x^2}}.$$

c) We start by differentiating $\cos\left(x^2\right)$. Denoting by $h(x) = \cos(x)$ and $g(x) = x^2$ we see that

$$\cos\left(x^2\right) = h\left(g(x)\right)$$

and as such, using the chain rule, we find that

$$\left(\cos\left(x^2\right)\right)' = h'(g(x))g'(x) = -\sin\left(x^2\right)\cdot 2x = -2x\sin\left(x^2\right).$$

Using the product rule of we find that

$$f'(x) = -2x\cdot\cos\left(x^2\right)+\left(2-x^2\right)\cdot\left(-2x\sin\left(x^2\right)\right) = 2x^3\sin\left(x^2\right)-4x\sin\left(x^2\right)-2x\cos\left(x^2\right).$$

d) Denoting by $h(x) = \sin(x)$ and $g(x) = \sin(x)$ we see that

$$f(x) = h\big(g(x)\big)$$

and as such, using the chain rule, we find that

$$f'(x) = h'(g(x))g'(x) = \cos(\sin(x))\cos(x).$$

e) We start by differentiating $\sqrt{4-x^2}$. Denoting by $h(x) = \sqrt{x} = x^{\frac{1}{2}}$ and $g(x) = 4 - x^2$ we see that

$$\sqrt{4-x^2} = h\big(g(x)\big)$$

and as such, using the chain rule, we find that

$$\left(\sqrt{4-x^2}\right)' = h'(g(x))g'(x) = \frac{1}{2}\left(4-x^2\right)^{-\frac{1}{2}} \cdot (-2x) = -\frac{x}{\sqrt{4-x^2}}.$$

Using the quotient rule we find that

$$f'(x) = \frac{1 \cdot \sqrt{4-x^2} - x \cdot \left(-\frac{x}{\sqrt{4-x^2}}\right)}{4-x^2} = \frac{\left(4-x^2\right) + x^2}{\sqrt{4-x^2}\left(4-x^2\right)} = \frac{4}{\left(4-x^2\right)^{\frac{3}{2}}}.$$

□

Example 3. Find the derivatives of the following functions:

a) $f(x) = e^{3x-1}$.
b) $f(x) = e^{4x^2}$.
c) $f(x) = e^{\sqrt{x}}$.
d) $f(x) = \ln(1+x^2)$.
e) $f(x) = \ln\left(\sqrt{1+x^2}\right)$.
f) $f(x) = \ln(\ln(x))$.

a) Using the chain rule we have that

$$f'(x) = e^{3x-1}(3x-1)' = 3e^{3x-1}.$$

b) Using the chain rule we have that

$$f'(x) = e^{4x^2}\left(4x^2\right)' = 8xe^{4x^2}.$$

c) Using the chain rule we have that

$$f'(x) = e^{\sqrt{x}}\left(\sqrt{x}\right)' = \frac{e^{\sqrt{x}}}{2\sqrt{x}}.$$

d) Using the chain rule we have that

$$f'(x) = \frac{1}{1+x^2} \cdot \left(1+x^2\right)' = \frac{2x}{1+x^2}.$$

e) We can use the chain rule with the functions $h(x) = \ln(x)$ and $g(x) = \sqrt{1+x^2}$ but remembering that
$$\ln\left(a^b\right) = b \ln a$$
we see that
$$f(x) = \ln\left(\left(1+x^2\right)^{\frac{1}{2}}\right) = \frac{\ln\left(1+x^2\right)}{2}.$$
Thus, using the previous part of the question and the linearity of the derivative we find that
$$f'(x) = \frac{1}{2} \cdot \frac{2x}{1+x^2} = \frac{x}{1+x^2}.$$

f) Using the chain rule we have that
$$f'(x) = \frac{1}{\ln(x)} (\ln(x))' = \frac{1}{\ln(x)} \cdot \frac{1}{x} = \frac{1}{x \ln(x)}.$$

2. Optimisation

The derivative is an essential tool in investigating the local, and global, minimum and maximum of a given function, defined as follows:

Definition 2. Let f be a given function on a interval (possibly infinite) I. We say that x_0 is a local minimum of f if there exists a neighbourhood of x_0, $U \subset I$ where
$$f(x_0) \leq f(x) \quad \text{for all } x \in U.$$
x_0 is called a global minimum of f on I if
$$f(x_0) \leq f(x) \quad \text{for all } x \in I.$$
Similarly, one can define a local and global maximum by demanding that $f(x_0) \geq f(x)$ in an appropriate neighbourhood U (local) or the entire interval I (global).

The key connections between the derivative and the investigation of the etrema of the function is given in the following theorem:

Theorem 2.1. *Let $f : I \to \mathbb{R}$ be differentiable. Then*

(i) If x_0 is an local etremum (minimum or maximum) then $f'(x_0) = 0$.
(ii) If $f'(x) > 0$ in an interval (a,b) then f is strictly increasing in (a,b), i.e. if $x_1 > x_2$ then $f(x_1) > f(x_2)$.
(iii) If $f'(x) \geq 0$ in an interval (a,b) then f is increasing (or non-decreasing) in (a,b), i.e. if $x_1 > x_2$ then $f(x_1) \geq f(x_2)$.
(iv) If $f'(x) < 0$ in an interval (a,b) then f is strictly decreasing in (a,b), i.e. if $x_1 > x_2$ then $f(x_1) < f(x_2)$.
(v) If $f'(x) \leq 0$ in an interval (a,b) then f is decreasing (or non-increasing) in (a,b), i.e. if $x_1 > x_2$ then $f(x_1) \leq f(x_2)$.

Corollary 2.2. When searching for a local extrema for a differentiable function f, we must solve the equation $f'(x) = 0$.

How do we search for a local extrema?

Step 1: Find all the candidates for local extrema by solving the equation $f'(x) = 0$.
Step 2: Draw the interval I as a line, and add to it the points you've found in the previous step and points where one can't differentiate (we won't see such points in out module).
Step 3: Choose a random point in between two points on the plotted line and check the sign of the derivative in it. Mark + or − on the line to indicate it sign. A + sign implies that the function is strictly increasing in the interval between these points, while a − sign implies that the function is strictly decreasing between them.
Step 4: A point is a local minimum if the sign on the plotted line switches from − to +, and a local maximum if the sign on the plotted line switches from + to −.

The study of global extrema is a bit more involved, as we may need to consider the boundaries of the interval I. The process is as follows:

- Find all local extrema.
- Add the boundary points of the interval that are in the interval.
- Compare the values of the function on the aforementioned points, and the limit of the functions on the boundaries of the interval that are *not* in the interval (including infinity). If the minimal or maximal of these values is attained in a local extrema or a boundary point that is in the interval - it is a global extrema. Else, there is no global minimum or maximum, respectively.

Remark 2. A function that is differentiable on $[a, b]$, where $a < b$ are finite real numbers, will always have a global minimum and maximum. This is a corollary of an important theorem in Calculus. As such you know that in order to find global extrema on a closed bounded interval all you need to do is to

- Solve the equation $f'(x) = 0$.
- Compare the values of f on these points to $f(a)$ and $f(b)$, and find the points that give the maximum and minimum.

Example 4. Consider the function $f(x) = x^4 - 2x^2$.

a) Find all local minimum and maximum of f on $\mathbb{R}$.
b) Find the global minimum and maximum of f on $[-2, 2]$.

Solution. We won't draw the points on a line in these notes, but instead will investigate the appropriate intervals between the suspected points.

a) We start by solving the equation $f'(x) = 0$. Since
$$f'(x) = 4x^3 - 4x = 4x\left(x^2 - 1\right)$$
we see that
$$f'(x) = 0 \quad \Leftrightarrow \quad x = 0, \text{ or } x^2 = 1.$$
Thus, our candidates for local extrema are $x = 0, -1, 1$.
Let us investigate the intervals between these points:
   - $x < -1$: For $x = -2$ we have that $f'(-2) = -8(4-1) < 0$. The associated sign is $-$, and we know that f is strictly decreasing in this domain.
   - $-1 < x < 0$: For $x = -\frac{1}{2}$ we have that $f'\left(-\frac{1}{2}\right) = -2\left(\frac{1}{4} - 1\right) > 0$. The associated sign is $+$, and we know that f is strictly increasing in this domain.
   - $0 < x < 1$: For $x = \frac{1}{2}$ we have that $f'\left(\frac{1}{2}\right) = 2\left(\frac{1}{4} - 1\right) < 0$. The associated sign is $-$, and we know that f is strictly decreasing in this domain.
   - $x > 1$: For $x = 2$ we have that $f'(2) = 8(4-1) > 0$. The associated sign is $+$, and we know that f is strictly increasing in this domain.

   From the above we conclude that $x = -1$ and $x = 1$ are a local minimums (the sign changes from $-$ to $+$) and $x = 0$ is a local maximum (the sign changes from $+$ to $-$).
b) To find the global extrema in the interval we compare the values of f on $0, -1$ and 1 with the values of f on the boundary points -2 and 2.
$$f(-2) = 8, \quad f(-1) = -1, \quad f(0) = 0, \quad f(1) = -1, \quad f(2) = 8.$$
Thus our global maximum is attained at $x = -2$ and $x = 2$, and equals 8, and our global minimum is attained at $x = -1$ and $x = 1$ and equals -1.

□

Example 5. A group of miners are trapped underground at a depth of 300 metres. A rescue team starts at the bottom of an abandoned mine shaft that is 600 metres West of the trapped miners and has a depth of 100 metres. The rescue team decides to start digging horizontally towards the East, and then to dig directly towards the trapped miners (potentially in diagonally). At a depth of 100 metres the rock is soft and it takes only 5 minutes to dig one horizontal metre. However, at any depth below this, the rock is hard and it takes 13 minutes to dig a distance of one metre. Calculate the minimal number of hours that it takes to tunnel to the trapped miners.

Solution. Let us denote by x the length of the horizontal tunnel the rescue team has dug to the East before starting to dig directly towards the miners. x must be in $[0,600]$. Once the rescue team has dug this horizontal tunnel, they are located $600-x$ metres West of the miners, and 200 metres above them. The distance between the team and the miners is thus $\sqrt{(600-x)^2+200^2}$. As the time it took the rescue team to dig the first stretch is $5x$ minutes (soft rock), and the time it took it to dig the second stretch is $13\sqrt{(600-x)^2+200^2}$ minutes (hard rock) we see that the total time in minutes the team must dig is

$$T(x) = 5x + 13\sqrt{(600-x)^2+200^2}.$$

We want to find the minimum of $T(x)$ over $[0,600]$. Differentiating $T(x)$ we see that

$$T'(x) = 5 + \frac{13}{2\sqrt{(600-x)^2+200^2}} \cdot (-2(600-x)).$$

The equation $T'(x) = 0$ is equivalent to

$$\frac{13(600-x)}{\sqrt{(600-x)^2+200^2}} = 5$$

or

$$13(600-x) = 5\sqrt{\left(600-x^2\right)+200^2}.$$

Squaring the above we get that

$$169(600-x)^2 = 25(600-x)^2 + 25\cdot 200^2$$

or $144(600-x)^2 = 5^2\cdot 200^2$. Since $600-x \geq 0$ we conclude that

$$(600-x) = \sqrt{\frac{5^2\cdot 200^2}{12^2}} = \frac{5\cdot 200}{12} = \frac{250}{3}.$$

Our potential minimums are $x=0$, $x = 600 - \frac{250}{3} = \frac{1550}{3}$ and $x = 600$ As

$$T(600) = 5600, \quad T\left(\frac{1550}{3}\right) = 5400, \quad T(0) = 2600\cdot\sqrt{10} > 5400,$$

we see that the minimal time the rescue team will dig is 5400 minutes, i.e. 90 hours. $\square$

3. Integration Part I - The Indefinite Integral (or Anti Derivative)

The indefinite integral is nothing more than the reverse of differentiation. Denoted by $\int f(x)dx$, the indefinite integral answers the question

Which function has f as its derivative?

A function whose derivative is f is called *a primitive function* of f. The indefinite integral of f, also known as the anti derivative of f, gives us the family of primitive functions of f. A natural question is: Is it well defined? The answer to this question lies with the next theorem

Theorem 3.1. *If F and G are primitive functions of f then there exists a constant $C \in \mathbb{R}$ such that $F = G + C$.*

The above theorem gives us the tools to find the anti derivative of f:

- Find a primitive function to f, F.
- The anti derivative of f, $\int f(x)dx$, equals to the family $\{F(x) + C \mid C \in \mathbb{R}\}$. We will write
$$\int f(x)dx = F(x) + C$$
to represent this family.

The above brings into light one big issue with the anti derivative: It is not unique. A more substantial issue with it is its complexity. While differentiation is quite straight forward, we can't compute explicitly most anti derivatives.
Much like differentiation, however, and in fact stemming from the appropriate rules for differentiation, the anti derivative does enjoy some properties that will help us compute it:

Theorem 3.2 (Rules of Integration)**.** *Let f and g be real valued functions. Then*

(i) (Linearity) For any $\alpha, \beta \in \mathbb{R}$
$$\int \left(\alpha f(x) + \beta g(x)\right) dx = \alpha \int f(x)dx + \beta \int g(x)dx.$$
(ii) (Integration by Parts {reverse product rule})
$$\int f'(x)g(x)dx = f(x)g(x) - \int f(x)g'(x)dx.$$
(iii) (u−substitution {reverse chain rule})
$$\int f\left(g(x)\right) g'(x)dx\Big|_{\text{at the point } x} = \int f(u)du\Big|_{\text{at the point } u=g(x)}.$$

We also have the following table to assist us

Table of Integration

Function	Anti Derivative
$f(x)=1$	$\int f(x)dx = x+C$
$f(x)=x^n$	$\int f(x)dx = \frac{x^{n+1}}{n+1}+C$
$f(x)=x^\alpha,\ \alpha\neq -1$	$\int f(x)dx = \frac{x^{\alpha+1}}{\alpha+1}+C$
$f(x)=\sin(x)$	$\int f(x)dx = -\cos(x)+C$
$f(x)=\cos(x)$	$\int f(x)dx = \sin(x)+C$
$f(x)=\frac{1}{\cos^2(x)}(x)$	$\int f(x)dx = \tan(x)+C$
$f(x)=-\frac{}{\sin^2(x)}$	$\int f(x)dx = \cot(x)+C$
$f(x)=e^x$	$\int f(x)dx = e^x+C$
$f(x)=a^x,\ a>0,\ a\neq 1$	$\int f(x)dx = \frac{a^x}{\ln(a)}+C$
$f(x)=\frac{1}{x}$	$\int f(x)dx = \ln\lvert x\rvert+C$

A few very useful formulas, stemming from the above and $u-$substitution, are

$$\int u(x)^\alpha u'(x)dx = \begin{cases} \frac{u(x)^{\alpha+1}}{\alpha+1}+C & \alpha\neq -1 \\ \ln|u(x)|+C & \alpha=-1 \end{cases}$$

$$\int \sin(u(x))\,u'(x)dx = -\cos(u(x))+C$$

$$\int \cos(u(x))\,u'(x)dx = \sin(u(x))+C$$

$$\int e^{u(x)}u'(x)dx = e^{u(x)}+C$$

$$\int a^{u(x)}u'(x)dx = \frac{a^{u(x)}}{\ln a}+C, \quad \text{if } a>0 \text{ and } a\neq 1.$$

These expressions are all particular cases of the following formula:

$$\int f'(u(x))\,u'(x)dx = f(u(x))+C.$$

Example 6. Compute the following integrals:

a) $\int \left(x^2+1\right)dx$.

b) $\int (x+1)\left(x^3-2\right)dx$.

c) $\int \left(1+\sqrt{x}\right)^2 dx$ (for $x>0$).

Solution. a) Due to linearity we see that

$$\int \left(x^2+1\right)dx = \int x^2 dx + \int 1\,dx = \frac{x^3}{3}+x+C.$$

b) Using linearity again we find that

$$\int (x+1)\left(x^3-2\right)dx = \int \left(x^4+x^3-2x-2\right)dx = \frac{x^5}{5}+\frac{x^4}{4}-x^2-2x+C.$$

c) Again, the linearity of the anti derivative implies that

$$\int \left(1+\sqrt{x}\right)^2 dx = \int \left(1+2\sqrt{x}+x\right) dx = x + 2\frac{x^{\frac{3}{2}}}{\frac{3}{2}} + \frac{x^2}{2} + C$$

$$= x + \frac{4x^{\frac{3}{2}}}{3} + \frac{x^2}{2} + C.$$

□

How does one use integration by parts?

- identify a function which we would like to differentiate to simplify our expression This function will play the role of $g(x)$. The remaining function will play the role of $f'(x)$.
- Find *any* choice for a primitive of f', f.
- Use the integration by parts formula and hope we get something we can solve[2].

Example 7. Find the following integrals using integration by parts:

a) $\int x \sin(x)\, dx$.
b) $\int xe^x dx$.
c) $\int \ln(x)\, dx$.
d) (*) $\int x \ln(x)\, dx$.

Solution. a) As the derivative of x is 1, we would like to define $g(x) = x$. Consequently $f'(x) = \sin(x)$. As a primitive to $f'(x)$ we choose $f(x) = -\cos(x)$, and naturally $g'(x) = 1$. Thus

$$\int x \sin(x)\, dx = x(-\cos(x)) - \int 1 \cdot (-\cos(x))\, dx$$

$$= -x\cos(x) + \int \cos(x) dx = -x\cos(x) + \sin(x) + C.$$

b) Similar to the previous problem, we choose $f'(x) = e^x$ and $g(x) = x$. Then, a choice of primitive to $f'(x)$ is $f(x) = e^x$ and since $g'(x) = 1$ we have that

$$\int xe^x dx = xe^x - \int 1 \cdot e^x dx = xe^x - \int e^x dx = xe^x - e^x + C.$$

c) This integration seems more complicated to compute with integration by parts as we only have one function in the integral! This can be rectified by noticing that

$$\ln(x) = 1 \cdot \ln(x) = (x)' \ln(x).$$

---

[2]There are cases, which we won't discuss in these notes, where using integration by parts results in a recursive formula.

This time our choice is motivated by the identification of the derivative. Defining $f'(x) = 1$ and $g(x) = \ln(x)$, we have that a possible primitive to $f'(x)$ is $f(x) = x$ and $g'(x) = \frac{1}{x}$. Thus

$$\int \ln(x)\, dx = x\ln(x) - \int x \cdot \frac{1}{x} dx = x\ln(x) - \int 1\, dx = x\ln(x) - x + C.$$

d) This time the function that we would like to differentiate is $g(x) = \ln(x)$ as $g'(x) = \frac{1}{x}$. Thus $f'(x) = x$ and a possible primitive to $f'(x)$ is $f(x) = \frac{x^2}{2}$.Thus

$$\int x\ln(x)\, dx = \frac{x^2}{2} \cdot \ln(x) - \int \frac{x^2}{2} \cdot \frac{1}{x} dx = \frac{x^2 \ln(x)}{2} - \frac{1}{2}\int x dx.$$

$$= \frac{x^2 \ln(x)}{2} - \frac{x^2}{4} + C.$$

□

How does one uses $u-$substitution?

- Identify a function that seems to "complicate" the integration, and name it $u(x)$.
- Find $u'(x)$ and using the symbolic notation $du = u'(x)dx$ get rid of the symbolic dx.
- Write the remaining functions with u - "eliminating" all the $x-$s.
- Integrate in u, and insert the expression for $u(x)$ in the final result.
- If possible *always* choose $u(x)$ in such a way that you can recognise $u'(x)dx$ in the integrand.

Example 8. Find

a) $\int \frac{\sin x}{\sqrt{\cos^3 x}} dx$.

b) $\int \frac{x^5}{\sqrt{1-x^6}} dx$.

c) $\int \frac{\sin(\sqrt{x+1})}{\sqrt{x+1}} dx$.

Solution. a) Recognising $\sin(x)$ as the almost derivative of $\cos(x)$ we define $u(x) = \cos(x)$. We have that $u'(x) = -\sin(x)dx$ and with the symbolic notation

$$du = -\sin(x)dx$$

we see that

$$\int \frac{\sin x}{\sqrt{\cos^3 x}} dx = \int \frac{-du}{\sqrt{u^3}} = -\int u^{-\frac{3}{2}} du = -\frac{u^{-\frac{1}{2}}}{-\frac{1}{2}}$$

$$= \frac{2}{\sqrt{u}} + C \underset{\text{bring } x \text{ back}}{=} \frac{2}{\sqrt{\cos(x)}} + C.$$

b) Recognising x^5 as almost the derivative of $1-x^6$ we define $u(x)=1-x^6$. When have that $u'(x)=-6x^5$ and with the symbolic notation

$$du = -6x^5 dx$$

we see that

$$\int \frac{x^5}{\sqrt{1-x^6}}dx = \int \frac{-\frac{du}{6}}{\sqrt{u}} = -\frac{1}{6}\int u^{-\frac{1}{2}}du = -\frac{1}{6}\frac{u^{\frac{1}{2}}}{\frac{1}{2}}$$

$$= -\frac{\sqrt{u}}{3} + C \underset{\text{bring } x \text{ back}}{=} -\frac{\sqrt{1-x^6}}{3} + C.$$

c) A derivative is not so easily recognisable here. Therefore, we will take the function that seems to "complicate" matters, $\sqrt{1+x}$, and call it $u(x)$. We have that

$$u'(x) = \frac{1}{2\sqrt{1+x}}$$

meaning that

$$du = \frac{1}{2\sqrt{1+x}}dx.$$

Noting that the right hand side appears in our integral, up to a factor of 2, we find that

$$\int \frac{\sin\left(\sqrt{x+1}\right)}{\sqrt{x+1}}dx = \int \sin(u)\cdot 2du = 2\int \sin(u)du = -2\cos(u)du + C$$

$$\underset{\text{bring } x \text{ back}}{=} -2\cos\left(\sqrt{1+x}\right) + C.$$

□

4. Integration Part II - The Definite Integral

The definite integral of a function f on the interval $[a,b]$, denoted by $\int_a^b f(x)dx$, answers the following question:

What is the area ,with allowed negative heights, "under" the graph of the function f over the interval $[a,b]$?

At first glance, the definite integral seems completely unrelated to the indefinite integral. A beautiful connection exists, however, and is given in the following theorem:

Theorem 4.1. *Let F be a primitive function of f over $[a,b]$. Then*

$$\int_a^b f(x)dx = F(b) - F(a).$$

It is worth to note, even though we won't use it in our module, that the above is a consequence of the following pivotal theorem:

Theorem 4.2 (The Fundamental Theorem of Calculus)**.** *Let f be an integral function on $[a,b]$. Then the function*

$$F(x)=\int_a^x f(t)dt$$

is a primitive function for f on (a,b).

Most of the anti derivative rules we've learned transfer seamlessly to the definite integral in the following way:

Theorem 4.3. *Let f and g be real valued functions. Then*

(i) (Linearity) For any $\alpha,\beta\in\mathbb{R}$

$$\int_a^b \left(\alpha f(x)+\beta g(x)\right)dx=\alpha\int_a^b f(x)dx+\beta\int_a^b g(x)dx.$$

(ii) (Integration by Parts {reverse product rule})

$$\int f'(x)g(x)dx=f(x)g(x)\Big|_a^b-\int_a^b f(x)g'(x)dx,$$

where $F(x)\Big|_a^b=F(b)-F(a)$.

(iii) ($u-$substitution {reverse chain rule})

$$\int_a^b f\left(g(x)\right)g'(x)dx=\int_{u(a)}^{u(b)} f(u)du.$$

Questions relating to definite integrals consist, mostly *but not always*, of finding a primitive function as we've done in the previous section.

Example 9. Compute the area beneath the function $f(x)=x^3-2x+4$ over the interval $[0,1]$.

Solution. The required area is $\int_0^1 f(x)dx$. We have that

$$\begin{aligned}\int_0^1 f(x)dx=\int_0^1\left(x^3-2x+4\right)dx&=\left(\frac{x^4}{4}-x^2+4x\right)\Big|_0^1\\&=\left(\frac{1}{4}-1+4\right)-\left(\frac{0}{4}-0+0\right)=\frac{17}{4}.\end{aligned}$$

□

Example 10. Compute the area beneath the function $f(x)=x\cos(x)$ over the interval $[0,\pi]$.

Solution. As the anti derivative is not immediate, we turn our attention to other methods. Since we have a multiplication of two functions, one of which we would be happy to differentiate (the function x), we consider using integration by parts. Setting $f'(x)=\cos(x)$ and $g(x)=x$, we have that $g'(x) = 1$ and a choice for a primitive function for $f'(x)$ is $f(x) = \sin(x)$. Thus

$$\int_0^{\pi} f(x)dx = \int_0^{\pi} x\cos(x)dx = x\sin(x)\Big|_0^{\pi} - \int_0^{\pi} 1\cdot\sin(x)dx$$

$$= \underbrace{(\pi\sin(\pi) - 0\cdot\sin(0))}_{=0} - \int_0^{\pi}\sin(x)dx = \cos(x)\Big|_0^{\pi} = \cos(\pi) - \cos(0) = -2.$$

□

Example 11. Compute the area beneath the function $f(x) = x^2\sin\left(x^3\right)$ over the interval $\left[0, \sqrt[3]{\frac{\pi}{2}}\right]$.

Solution. Recognising x^2 as the almost derivative of x^3 we find ourselves tempted to use $u-$substitution with $u(x) = x^3$. In that case we have that $u'(x) = 3x^2$, meaning that

$$du = 3x^2 dx.$$

Thus

$$\int_0^{\sqrt[3]{\frac{\pi}{2}}} f(x)dx = \int_0^{\sqrt[3]{\frac{\pi}{2}}} x^2\sin\left(x^3\right)dx = \int_{0^3}^{\left(\sqrt[3]{\frac{\pi}{2}}\right)^3}\sin(u)\frac{du}{3}$$

$$= \frac{1}{3}\int_0^{\frac{\pi}{2}}\sin(u)du = \frac{-\cos(u)}{3}\Big|_0^{\frac{\pi}{2}} = \frac{-\cos\left(\frac{\pi}{2}\right) - (-\cos(0))}{3} = \frac{1}{3}.$$

□