5. Derivatives and integrals

These lecture notes present my interpretation of Ruth Lawrence's lecture notes (in Hebrew)

[1]

5.1 The derivative of a function

5.1.1 Definition

A function defined on the real line, or on a segment, contains an infinite amount of information. Consider a function $f:\mathbb{R}\to\mathbb{R}$. We can evaluate it at a number of points,

$$f(a_1), f(a_2), \dots, f(a_n),$$

but this information does not reveal the values returned by f for other inputs.

A function represents a relation between two quantities—an input and an output. The domain may represent, for example, time in hours taking values between 0 and 24, and the range may represent temperature in Celsius assuming values greater than -273. Very often one is interested in knowing how the output changes upon a change in the input.

Take a function $f:A\to B$ and let $a\in A$ be a number in its domain. Let $x\in A$ be another point in the domain of f. The change in the value returned by f due to the input changing from a to x is

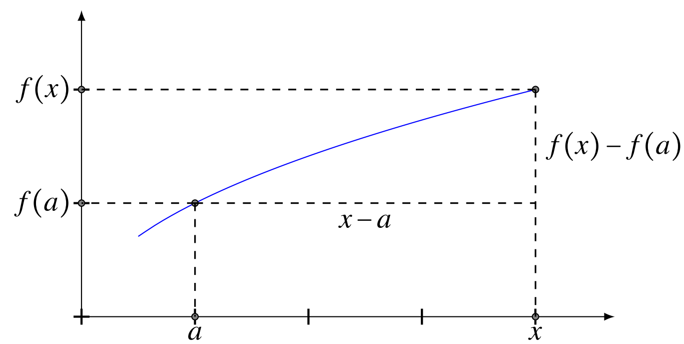
$$f(x)-f(a).$$

How can we quantify whether this change is large or small? This seems a meaningless question, until we realize that the change in the output can be measured with

[1] Image of Gottfried Wilhelm Leibniz, 1646–1716

respect to the change in the input. The ***mean rate of change*** (קצב השינוי הממוצע) in the value of f due to modifying the input from a to x is

$$\frac{f(x)-f(a)}{x-a}. \tag{5.1}$$



Comments 5.1

1. The ratio (5.1) is not defined for $x=a$.
2. If this ratio is positive then either $x>a$ and $f(x)>f(a)$, or $x<a$ and $f(x)<f(a)$. In either case, the output and the input change in the same direction.
3. If this ratio is negative then either $x>a$ and $f(x)<f(a)$, or $x<a$ and $f(x)>f(a)$. In either case, the output and the input change in opposite directions.
4. Suppose that x represents time in hours and $f(x)$ represents position in kms along a straight road. Then this ratio is what we call the ***mean velocity*** in km/h between time a and time x.

Consider the mean change of f between a and x. If we assume that f and a are given, then it is only a function of x, which we denote by

$$\Delta_{f,a}(x) = \frac{f(x)-f(a)}{x-a}. \tag{5.2}$$

Think now what happens as we change x and make it closer and closer to a. The dominator becomes smaller and smaller (in absolute value). The behavior of the numerator depends on the function f, but if f is continuous the numerator decreases as well (in absolute value). If the function $\Delta_{f,a}$ has a ***limit*** as x tends to a, this limit represent an ***instantaneous rate of range*** (קצב שינוי נקודתי) of f at a. If this limit exists (and note this important "if"), we call it the ***derivative (נגזרת) of*** f ***at*** a, and denote it by

$$f'(a) = \lim_{x\to a}\Delta_{f,a}(x) = \lim_{x\to a}\frac{f(x)-f(a)}{x-a}.$$

We also say that f is ***differentiable*** (גזירה) at a.

It is often more convenient to express x as $a+h$ (h is then the deviation from a), in which case

$$f'(a) = \lim_{h\to 0} \Delta_{f,a}(a+h) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}.$$

Comments 5.2

1. We haven't defined what we mean by "limit". We will leave this notion fuzzy for the moment and assume you understand it intuitively. You'll spend most of 80177 understanding what a limit is.
2. When we write $x \to a$ or $h \to 0$, the tendency is both from above and below; h may approach 0 from above, from below, or oscillate in sign.
3. Another common notation for the derivative of f at a is

$$\frac{df}{dx}(a).$$

   This notation, due to Leibniz, is less "clean" that $f'(a)$. What does x represent in this notation?

■ **Example 5.1** Let $f : x \mapsto x^2$ and $a = 3$. Then,

$$\Delta_{f,a}(3+h) = \frac{(3+h)^2 - 3^3}{h} = \frac{9+6h+h^2-9}{h} = 6+h.$$

As $h \to 0$ this expression has a limit—6, hence

$$f'(3) = \lim_{h\to 0} \Delta_{f,a}(3+h) = 6.$$

■

■ **Example 5.2** Let $f : x \mapsto |x|$ and $a = 0$. Then,

$$\Delta_{f,a}(a+h) = \frac{|0+h|-|0|}{h} = \frac{|h|}{h} = \begin{cases} 1 & h > 0 \\ -1 & h < 0. \end{cases}$$

This expression does not have a limit as $h \to 0$ (it only has ***one-sided limits***) hence f is not differentiable at zero,

$$f'(0) \quad \text{does not exist.}$$

■

■ **Example 5.3** Consider the Heaviside function

$$f : x \mapsto \begin{cases} 0 & x < 0 \\ 1 & x \ge 0. \end{cases}$$

Let $a = 2$. Then,

$$\Delta_{f,2}(2+h) = \begin{cases} -1/h & h < -2 \\ 0 & h \ge -2. \end{cases}$$

Since we're interested in the limit of this function as $h \to 0$ we may well ignore values of h less than -2. Close enough to $h = 0$, $\Delta_{f,a}(a+h)$ is identically zero, hence

$$f'(2) = \lim_{h\to 0} \Delta_{f,2}(2+h) = 0.$$

On the other hand, take $a = 0$. Then,

$$\Delta_{f,0}(0+h) = \begin{cases} -1/h & h < 0 \\ 0 & h \geq 0, \end{cases}$$

and the limit $h \to 0$ does not exist, i.e.,

$$f'(0) \qquad \text{does not exist.}$$

■

In fact we may state something more general:

Theorem 5.1 Let $f : \mathbb{R} \to \mathbb{R}$. If f is not continuous at a then it is neither differentiable at a. (Equivalently, if f is differentiable at a then it is necessarily also continuous at that point.)

Proof. It is always true that for $x \neq a$,

$$f(x) = f(a) + \frac{f(x) - f(a)}{x - a}(x - a) = f(a) + \Delta_{f,a}(x)(x - a).$$

If f is differentiable at a then $\Delta_{f,a}(x)$ tends to $f'(a)$ as $x \to a$, and since $(x - a) \to 0$ it follows that

$$\lim_{x\to a} f(x) = f(a),$$

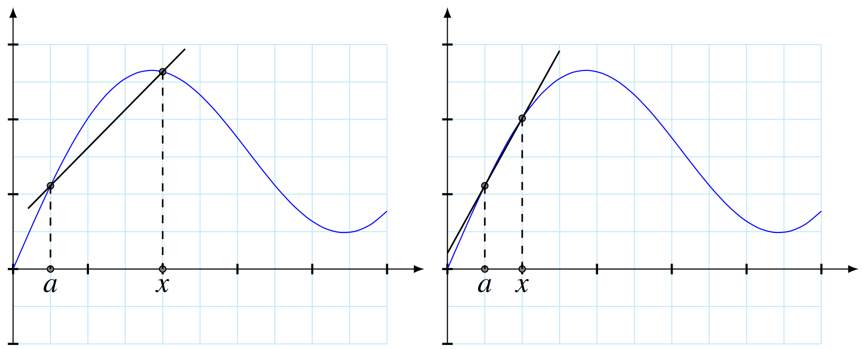
i.e., f is continuous at a. ■

5.1.2 Graphical representation

The derivative of a function at a point has an intuitive geometric interpretation based on the graph representation of a function. The ratio

$$\Delta_{f,a}(x) = \frac{f(x) - f(a)}{x - a}$$

is the slope of the ***chord*** (מיתר) that connects the points $(a, f(a))$ and $(x, f(x))$:



As $x \to a$ this chord converges to a line tangent (משיק) to the graph of f at a.

5.1.3 The derivative function

If a function $f : A \to B$ is differentiable on a subset $\tilde{A} \subset A$ of its domain, then we can define a new function $f' : \tilde{A} \to \mathbb{R}$—the ***derivative function*** (פונקצית הנגזרת). The derivative function can now be analyzed in its turn. For example, we may ask whether it is continuous, whether it is differentiable, and so on. The derivative of f' is called the ***second derivative*** of f, and it denoted f''. The derivative of f'' is denoted f''' or $f^{(3)}$. The n-th derivative of f (if it exists) is denoted $f^{(n)}$.

■ **Example 5.4** Consider the function $f : x \mapsto x^2$. For every $a \in \mathbb{R}$,

$$\Delta_{f,a}(a+h) = \frac{(a+h)^2 - a^2}{h} = 2a + h,$$

hence, letting $h \to 0$,

$$f'(a) = 2a,$$

i.e.,

$$f' : \mathbb{R} \to \mathbb{R}, \qquad f' : y \mapsto 2y.$$

To find the second derivative of f we set

$$\Delta_{f,'a}(a+h) = \frac{2(a+h) - 2a}{h} = 2,$$

and letting $h \to 0$,

$$f''(a) = 2,$$

i.e.,

$$f'' : \mathbb{R} \to \mathbb{R}, \qquad f'' : y \mapsto 2.$$

■

5.2 The laws of differentiation

Given a function $f : A \to B$, how do we find its derivative? We saw specific examples in which we found the derivative by applying the definition and calculating the limit of the mean rate of change. If our goal is to calculate the derivatives of a large collection of functions, this is not efficient. In this section we will review the laws of differentiation that enable us to calculate the derivatives of any function composed of functions whose derivatives are known.

5.2.1 Sums, products, ratios and compositions

Proposition 5.2 If the functions f and g are differentiable at a then so is the function $f+g$ and

$$(f+g)'(a) = f'(a) + g'(a).$$

Proof. Consider the identity,

$$\Delta_{f+g,a}(x) = \frac{(f(x)+g(x)) - (f(a)+g(a))}{x-a} = \Delta_{f,a}(x) + \Delta_{g,a}(x).$$

Letting $x \to a$ we recover the desired result. ■

Proposition 5.3 — Leibniz rule. If the functions f and g are differentiable at a then so is the function $f \cdot g$ and

$$(fg)'(a) = f'(a)g(a) + f(a)g'(a).$$

Proof. Consider the identity,

$$\begin{aligned}\Delta_{fg,a}(x) &= \frac{f(x)g(x) - f(a)g(a)}{x-a} \\ &= \frac{f(x)g(x) - f(a)g(x) + f(a)g(x) - f(a)g(a)}{x-a} \\ &= \Delta_{f,a}(x)g(x) + f(a)\Delta_{g,a}(x).\end{aligned}$$

Letting $x \to a$ we recover the desired result. ■

Proposition 5.4 If the function f is differentiable at a and $f(a) \neq 0$, then $1/f$ is differentiable at a and

$$(1/f)'(a) = -\frac{f'(a)}{f^2(a)}.$$

Proof. Consider the identity,

$$\Delta_{1/f,a}(x) = \frac{1/f(x) - 1/f(a)}{x-a} = \frac{f(a)-f(x)}{f(x)f(a)(x-a)} = -\Delta_{f,a}(x)\frac{1}{f(a)f(x)}.$$

Letting $x \to a$ we recover the desired result. ■

Corollary 5.5 If the functions f and g are differentiable at a and $g(a) \neq 0$, then the function f/g is differentiable at a and

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

Proof. Apply the last two rules,

$$(f/g)'(a) = (1/g)'(a)f(a) + \frac{1}{g(a)}f'(a) = -\frac{g'(a)}{g^2(a)}f(a) + \frac{1}{g(a)}f'(a).$$

■

Proposition 5.6 — Chain rule (כלל השרשרת). If f is differentiable at a and g is differentiable at $f(a)$ then $g \circ f$ is differentiable at a and

$$(g \circ f)'(a) = g'(f(a))f'(a).$$

Proof. Consider the identity,

$$\begin{aligned}\Delta_{g\circ f,a}(x) &= \frac{g(f(x)) - g(f(a))}{x-a} \\ &= \frac{g(f(x)) - g(f(a))}{f(x)-f(a)} \cdot \frac{f(x)-f(a)}{x-a} \\ &= -\Delta_{g,f(a)}(f(x))\Delta_{f,a}(x).\end{aligned}$$

Letting $x \to a$, we also have $f(x) \to f(a)$, since f is continuous. Thus the first term on the right-hand side converges to $g'(f(a))$, whereas the second term converges to $f'(a)$. ■

Thus, we can create a list of functions whose derivatives we find directly from the definition, and then enlarge this collection using the laws of differentiation. Below is a list of elementary functions whose derivative we must know by heart:

$f(x)$	$f'(x)$
x^n	nx^{n-1}
e^x	e^x
$\ln x$	$1/x$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$

■ **Example 5.5**

$$\sec' x = -\frac{\cos' x}{\cos^2 x} = \frac{\sin x}{\cos^2 x}.$$

■

■ **Example 5.6**

$$\tan' x = \frac{\cos x \sin' x - \sin x \cos' x}{\cos^2 x} = \frac{1}{\cos^2 x}.$$

■

■ **Example 5.7** Consider the function

$$f(x) = x^2 \sin x.$$

By the Leibniz rule,

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

Note that it is very common to write

$$\frac{d}{dx}\left(x^2 \sin x\right) = \frac{d}{dx}x^2 \sin x + x^2 \frac{d}{dx}\sin x = \text{etc.}$$

■

■ **Example 5.8** Consider the function

$$h(x) = \sin(x^2).$$

It is the composition $g \circ f$ of $g = \sin$ and $f : x \mapsto x^2$. Using the chain rule

$$h'(x) = g'(f(x))\, f'(x) = \cos(x^2)\, 2x.$$

■

■ **Example 5.9** Consider the function

$$h(x) = \sin(x \ln x).$$

It is the composition $g \circ f$ of $g = \sin$ and $f : x \mapsto x \ln x$. By the chain rule,

$$h'(x) = \sin'(f(x))\, f'(x) = \cos(x \ln x) \cdot \left(x \cdot \frac{1}{x} + \ln x\right).$$

■

■ **Example 5.10** Consider the function

$$\cosh x = \frac{e^x + e^{-x}}{2}.$$

Its derivative is given by

$$\cosh' x = \frac{e^x - e^{-x}}{2} = \sinh x.$$

Similarly, you may find that

$$\sinh' x = \cosh x.$$

■

5.2.2 The derivative of inverse functions

What is the derivative of arcsin? Does knowing the derivative of sin help up in finding the derivative of its inverse? Recall that the inverse f^{-1} of a function f is defined by

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

This is an identity between two functions, which we may also write as

$$f^{-1} \circ f = \text{id}.$$

Differentiating both sides, using the chain rule on the left-hand side,

$$f^{-1\prime}(f(x))\, f'(x) = 1,$$

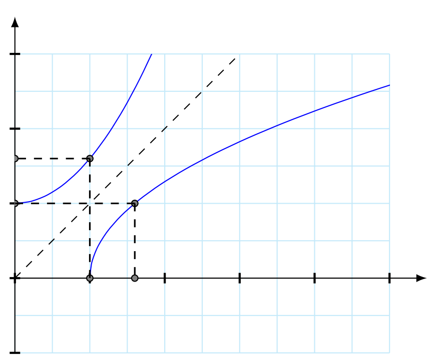
i.e.,

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

Substituting $f(x) = u$, i.e., $x = f^{-1}(u)$ we get the following formula for the derivative of an inverse function,

$$\boxed{f^{-1\prime}(u) = \frac{1}{f'(f^{-1}(u))}} \tag{5.3}$$

This formula is easy to understand by considering the graphs of a function and its inverse:



■ **Example 5.11** Let $f = \sin : [-\pi/2, \pi/2] \to [-1, 1]$ then by (5.3)

$$\arcsin'(u) = \frac{1}{\sin'(\arcsin(u))} = \frac{1}{\cos(\arcsin(u))}.$$

By the properties of the trigonometric functions,

$$\cos(\arcsin(u)) = \sqrt{1 - \sin^2(\arcsin(u))} = \sqrt{1 - u^2},$$

hence,

$$\arcsin'(x) = \frac{1}{\sqrt{1 - x^2}}.$$

■

■ **Example 5.12** Suppose the only information we had on the function $\exp : \mathbb{R} \to (0, \infty)$ was that $\exp' = \exp$. Denote its inverse by $\ln = \exp^{-1}$. Then, by (5.3)

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

■

5.2.3 The derivative of implicit functions

Recall that a function $A \to B$ may be defined by means of its graph, i.e., a set of points $(x, y) \in A \times B$. So far we only considered functions defined ***explicitly*** (מפורשות),

$$\text{Graph}(f) = \{(x, y) \mid y = f(x)\},$$

which means that we have an expression that tells us the value of y given x. Functions can also be defined ***implicitly*** (באופן סתום) as in the following example,

■ **Example 5.13** Consider a function f given by the following graph,

$$\text{Graph}(f) = \{(x, y) \mid xy^3 + x^3 y = 10\}.$$

In what sense this is a function? To every x corresponds a y, which we denote by $f(x)$, satisfying that $(x, f(x)) \in \text{Graph}(f)$. Thus,

$$x f^3(x) + x^3 f(x) = 10. \tag{5.4}$$

What is the derivative of f at x? Equation (5.4) is an identity between two functions,

$$G(x) = H(x),$$

where

$$G(x) = x f^3(x) + x^3 f(x) \qquad \text{and} \qquad H(x) = 10.$$

Their derivatives must also be equal,

$$G'(x) = H'(x),$$

namely,

$$f^3(x) + 3x f^2(x) f'(x) + 3x^2 f(x) + x^3 f'(x) = 0,$$

or,

$$f'(x) = -\frac{f^3(x) + 3x^2 f(x)}{3x f^2(x) + x^3}.$$

Given x, if we know $f(x)$ then we can evaluate $f'(x)$ as well. For example, you may verify that $(1,2) \in \text{Graph}(f)$, i.e., $f(1) = 2$, and therefore

$$f'(x) = -\frac{14}{13}.$$

■

5.3 Linear approximation

Definition 5.1 Let V and W be vector spaces. A function $f : V \to W$ is called ***linear*** if for all $x, y \in V$ and $\alpha, \beta \in \mathbb{R}$,

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y).$$

The space of linear functions $V \to W$ is denoted by $L(V,W)$.

The real number are a vector space hence function $f : \mathbb{R} \to \mathbb{R}$ are functions between two vector spaces. The space of linear functions $L(\mathbb{R},\mathbb{R})$ consists of all the functions that multiply their input by a constant number,

$$x \mapsto ax.$$

How do linear function relate to derivatives? The value of a function f at a point a is a crude information about $f(x)$ in the vicinity of a. If f is continuous at a then when x is “very close" to a, $f(x)$ also is “very close" to $f(a)$,

$$f(x) \approx f(a) \qquad \text{when } x \approx a,$$

which is a quantitatively meaningless statement. If we know in addition the derivative of f at a, then we know that

$$\frac{f(x) - f(a)}{x - a} \approx f'(a) \qquad \text{when } x \approx a,$$

i.e.,

$$f(x) - f(a) \approx f'(x)(x - a) \qquad \text{when } x \approx a. \tag{5.5}$$

Again, we have no control on what $\approx$ means.

Equation (5.5) states that when x is close to a, the deviation of $f(x)$ from $f(a)$ is approximately a linear function of $(x - a)$ (the constant of proportionality is precisely the derivative of f at a). The approximation of a function using its value and the value of its derivative at a point is called a ***linear approximation***. Graphically, it amounts to approximating the graph of a function by its tangent line at a.

■ **Example 5.14** The linear approximation of ln using its value and the value of its derivative at 1 is

$$\ln x \approx \ln 1 + \frac{1}{1} \cdot (x-1),$$

i.e.,

$$\ln x \approx x - 1,$$

or,

$$\ln(1+x) \approx x.$$

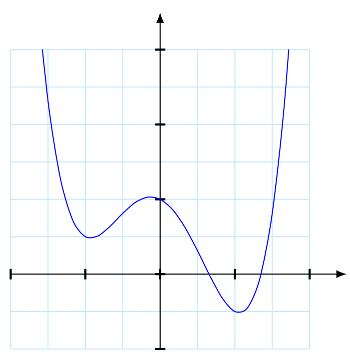
■

5.4 Derivative and extrema

Recall that a continuous function defined on a closed interval $[a,b]$ has both minima and maxima in that interval. These extrema may be the points a,b or ***internal points*** (נקודות פנים).

Definition 5.2 A point x is called a ***local minimum*** of f is there exists a neighborhood of x in which f assume its minimum at x. Similarly, x is called a ***local maximum*** of f is there exists a neighborhood of x in which f assume its maximum at x

The figure below shows a function that has two local minima, one of which is a global minimum, and one local maximum, which is not a global maximum.



Theorem 5.7 — Fermat. If a is a local maximum (resp. minimum) of f and f is differentiable at a, then $f'(a) = 0$.

Proof. Since there exists a neighborhood of a in which

$$\text{for all } x \qquad f(a) \ge f(x),$$

it follows that

$$\Delta_{f,a}(a+h) = \frac{f(a+h)-f(a)}{h} \begin{cases} \le 0 & h > 0 \\ \ge 0 & h < 0. \end{cases}$$

Since f is differentiable at a, the limit of $\Delta_{f,a}(a+h)$ as $h \to 0$ exists. Because the sign of $\Delta_{f,a}(a+h)$ is non-negative on one side and non-positive on the other side, the limit must be zero. ■

Comment 5.1 Fermat's theorem does not state that $f'(a) = 0$ implies that f is a local extremum.

5.5 The mean-value theorem

Theorem 5.8 — Rolle. Let $f : [a,b] \to \mathbb{R}$ be continuous in $[a,b]$ and differentiable in (a,b), and

$$f(a) = f(b),$$

then there exists a point $c \in (a,b)$ such that

$$f'(c) = 0.$$

Proof. We know that f must have both minima and maxima. If both are end points then f is constant and its derivative vanishes identically. If, otherwise, there is an internal point c which is extremal, then by Fermat's theorem $f'(c) = 0$. ■

Theorem 5.9 — mean-value (הערך הממוצע). Let $f : [a,b] \to \mathbb{R}$ be continuous in $[a,b]$ and differentiable in (a,b), then there exists a point $c \in (a,b)$ such that

$$f'(c) = \frac{f(b)-f(a)}{b-a}.$$

Proof. Consider the function

$$h(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a),$$

and apply Rolle's theorem. ∎

Comment 5.2 The mean-value theorem implies that if you drive 60 km in 1 hour, then at some time your speed must have been 60 km/h.

Corollary 5.10 If $f : [a,b] \to \mathbb{R}$ is continuous in $[a,b]$ and differentiable in (a,b) and furthermore

$$\text{for all } x \in (a,b) \qquad f'(x) > 0,$$

then f is (strictly) monotonically increasing in $[a,b]$.

Proof. If $c < d$ are in $[a,b]$, then there exists a point e such that

$$\frac{f(d) - f(c)}{d - c} = f'(e) > 0,$$

from which follows that $f(d) > f(c)$. ∎

Corollary 5.11 If $f : [a,b] \to \mathbb{R}$ is continuous in $[a,b]$ and differentiable in (a,b) and furthermore

$$\text{for all } x \in (a,b) \qquad f'(x) = 0,$$

then f constant in $[a,b]$.

Proof. For every x in (a,b) there exists a point e such that

$$\frac{f(x) - f(a)}{x - a} = f'(e) = 0,$$

from which follows that $f(x) = f(a)$. ∎

Comment 5.3 Proving that the derivative of a constant function is zero is straightforward. Proving the other direction requires the mean-value theorem.

5.6 The definite integral

5.6.1 Definition of the integral

Analytic geometry provides us with a tool for reasoning about geometry using analytic methods. In particular, it provides us with tools for dealing with the notion of ***area*** (which is a measure of magnitude of a two-dimensional figure).

Like in any measurement, when it comes to measuring area one need a ***unit of measurement*** (יחידת מידה). The standard choice is a square of unit side length. The area of a figure is then the number of such unit squares needed to cover that figure (of course, a unit square can be partitioned into smaller parts, like any other unit of measurement).

Let f be a “nice enough" function (we’ll be slightly more precise below). Suppose for the moment that it is non-negative. Let $a,b\in\mathbb{R}$, $a<b$. The area between the graph of the function, the x-axis, and the vertical lines $x=a$ and $x=b$ is called the ***integral of*** f ***between*** a ***and*** b. I will denote it by

$$\int_a^b f.$$

The more standard notation is

$$\int_a^b f(x)\,dx \quad\text{or}\quad \int_a^b f(t)\,dt \quad\text{or}\quad \int_a^b f(\xi)\,d\xi.$$

I will comment about it below.

What is actually the meaning of the area when the shape is not a rectangle? Let’s start by considering a ***piecewise constant function*** (פונקציה קבועה למקוטעין). Let

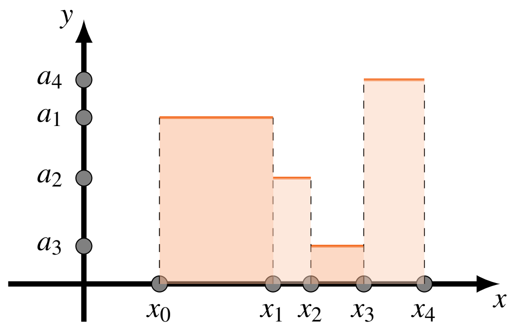
$$a=x_0<x_1<x_2<\cdots<x_n=b$$

be $n+1$ points that ***partition*** the segment $[a,b]$ into n sub-segments, and consider a function of the form

$$f(x)=\begin{cases} a_1 & x_0\le x<x_1\\ a_2 & x_1\le x<x_2\\ \vdots & \vdots\\ a_n & x_{n-1}\le x\le x_n.\end{cases}$$

Since the area under the graph of f is a union of rectangles it is given by

$$\int_a^b f=a_1(x_1-x_0)+a_2(x_2-x_1)+\cdots+a_n(x_n-x_{n-1})=\sum_{k=1}^n a_k(x_k-x_{k-1}).$$



Take now a general (nice) function. We can partition the interval $[a,b]$ into sub-intervals by selecting a set of points

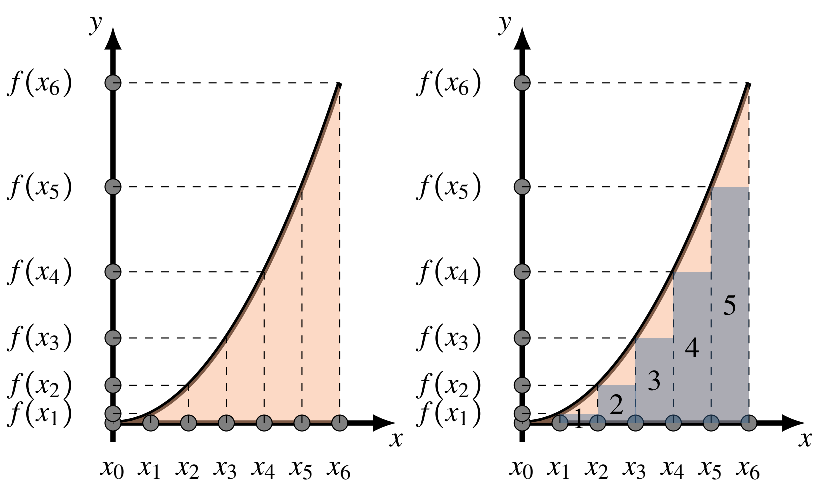
$$a = x_0 < x_1 < x_2 < \cdots < x_n = b,$$

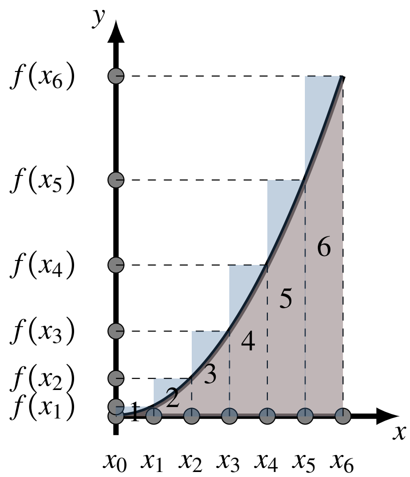
which may, but don't have to be equi-distanced.

In each sub-interval $[x_{k-1}, x_k]$ denote by m_k the minimum of the function and by M_k its maximum (they are guaranteed to exist only if the function is continuous and the interval is closed; otherwise there is a close notion of ***infimum*** and ***supremum***). It is guaranteed that

$$\sum_{k=1}^{n} m_k(x_k - x_{k-1}) \le \int_a^b f \le \sum_{k=1}^{n} M_k(x_k - x_{k-1}).$$

If, as the partition becomes finer and finer, the two sides of this inequality tend to each other (and both tend to a limit), we say that f is ***integrable*** and identify this limit as the integral of f between a and b (in fact, when we say that a function is "nice" in this context we mean that it is integrable).





■ **Example 5.15** Archimedes' method for calculating the area under a parabola. ■

5.6.2 The fundamental theorem of calculus

Having defined the notions of area and integral, there remains a practical question: how to calculate the area under the graph of a given function f?

First note that by the very definition of the integral as an area, for $a < b < c$ we have

$$\int_a^b f + \int_b^c f = \int_a^c f.$$

Given a function f and an interval (in its domain) $[a,b]$, define a new function, $F : [a,b] \to \mathbb{R}$, by

$$F(x) = \int_a^x f.$$

What can we say about this new function F?

1. $F(a) = 0$.
2. The mean rate of change of F between x and $x+h$ is

$$\Delta_{F,x}(x+h) = \frac{F(x+h) - F(x)}{h} = \frac{1}{h}\left(\int_a^{x+h} f - \int_a^x f\right) = \frac{1}{h}\int_x^{x+h} f.$$

As h tends to zero, assuming that f is continuous at x, the integral becomes close to $h f(x)$, hence

$$F'(x) = \lim_{h\to 0} \Delta_{F,x}(x+h) = f(x).$$

Thus, assuming that f is continuous, the function F satisfies the following two properties:

$$F' = f \qquad \text{and} \qquad F(a) = 0.$$

The first property is that F is an ***anti-derivative*** of f (a function whose derivative is f, more commonly known as its ***primitive function*** (פונקציה קדומה)). Recall that if the anti-derivative of f is F then so is the derivative of any function of the form $F + C$. The anti-derivative we're looking for is one that vanishes at a.

Suppose that G is *some* anti-derivative of f, i.e., $G' = f$. Then,

$$F(x) = G(x) - G(a)$$

is the function we are looking for. In particular,

$$\boxed{\int_a^b f = F(b) = G(b) - G(a)} \tag{5.6}$$

The relation between the integral and the derivative is known as the ***fundamental theorem of calculus***. Formula (5.6) that expresses the integral of a function f between two points a and b in terms of *some* anti-derivative G, is known as the ***Newton-Leibniz formula***.

■ **Example 5.16** Let $f(x) = x^2$. What is

$$\int_2^5 f\,?$$

The function $G(x) = x^3/3 + 19$ is an anti-derivative of f. Thus,

$$\int_2^5 f = G(5) - G(2) = (125/3 + 19) - (8/3 + 19) = 39.$$

■

Comment 5.4 For a function f and points a, b we denote

$$f|_a^b = f(b) - f(a).$$

Thus the Newton-Leibniz formula takes the form

$$\int_a^b f = G|_a^b.$$

Finding anti-derivatives (also called ***indefinite integrals***) turns out to be a very useful thing. It is however much harder than finding a derivative, and we will devote a whole chapter for that purpose.

Comment 5.5 A few words about standard notations: the derivative of f between a and b is approximated by an expression of the type

$$\sum_{k=1}^{n} f(x_k)(x_k - x_{k-1}) \equiv \sum_{k=1}^{n} f(x_k)\,\Delta x_k,$$

where x_k is a point in the k-th interval and Δx_k is that width of that interval; such sum are known as ***Riemann sums***. As the partition becomes denser and denser the sum contains more and more terms. The integral sign $\int$ represents the summation sign $\sum$ as the number of summands tends to infinity. The symbol dx represents Δx_k as it gets smaller and smaller.

5.6.3 More about integrals and their derivatives

We have defined the integral

$$\int_a^b f$$

assuming that $a < b$. For $a < b < c$ we have

$$\int_a^b f = \int_b^c f = \int_a^c f. \tag{5.7}$$

If we define for $a < b$

$$\int_b^a f = -\int_a^b,$$

then the summation rule (5.7) remains valid for any a, b, c.

Let f be integrable and let g be differentiable. Consider the function

$$G(x) = \int_a^{g(x)} f,$$

which you're more likely to see written as

$$G(x) = \int_a^{g(x)} f(t)\,dt.$$

What is the derivative of G? If we denote as before

$$F(x) = \int_a^x f,$$

then $G = F \circ g$, hence by the chain rule,

$$G'(x) = F'(g(x))\,g'(x) = f(g(x))\,g'(x).$$

Similarly, if h is some other differentiable function and

$$H(x) = \int_{h(x)}^b f = -\int_b^{h(x)} f,$$

then

$$H'(x) = -f(h(x))\,h'(x).$$

Finally, if

$$J(x) = \int_{h(x)}^{g(x)} f,$$

then as for every constant a,

$$J(x) = \int_{h(x)}^{a} f + \int_{a}^{g(x)} f,$$

it follows that

$$J'(x) = f(g(x))\,g'(x) - f(h(x))\,h'(x).$$

Example 5.17 Consider the function

$$J(x) = \int_{\ln x}^{\sqrt{x}} \sin.$$

Then,

$$J'(x) = \sin(\sqrt{x}) \cdot \frac{1}{2\sqrt{x}} - \sin(\ln x) \cdot \frac{1}{x}.$$

■