Real Functions in One Variable

Leif Mejlbro

bookboon
The eBook company

Leif Mejlbro

Real Functions of One Variable

Calculus 1a

Real Functions of One Variable – Calculus 1a

ISBN 978-87-7681-117-4

Contents

Preface

The publisher recently asked me to write an overview of the most common subjects in a first course of Calculus at university level. I have been very pleased by this request, although the task has been far from easy.

Since most students already have their recommended textbook, I decided instead to write this contribution in a totally different style, not bothering too much with rigoristic assumptions and proofs. The purpose was to explain the main ideas and to give some warnings at places where students traditionally make errors.

By rereading traditional textbooks from the first course of Calculus I realized that since I was not bound to a strict logical structure of the contents, always thinking of the students' ability at that particular stage of the text, I could give some additional results which may be useful for the reader. These extra results cannot be given in normal textbooks without violating their general idea. This has actually been great fun to me, and I hope that the reader will find these additions useful. At the same time most of the usual stuff in these initial courses in Calculus has been treated.

When emphasizing formulæ I had the choice of putting them into a box, or just give them a number. I have chose the latter, because too many boxes would overwhelm the reader. On the other hand, I had sometimes also to number less important formulæ because there are local references to them. I hope that the reader can distinguish between these two applications of the numbering.

In the Appendix I have collected some useful formulæ, which the reader may use for references.

It should be emphasized that this is *not* an ordinary textbook, but instead a supplement to existing ones, hopefully giving some new ideas in how problems in Calculus can be solved.

It is impossible to avoid errors in any book, so even if I have done my best to correct them, I would not dare to claim that I have got rid of all of them. If the reader unfortunately should use a formula or result which has been wrongly put here (misprint or something missing) I do hope that my sins will be forgiven.

Leif Mejlbro

In the revision of these notes the publisher and I agreed that the title should be *Calculus* followed by a number (1–3) and a letter, where

a stands for *compendia*,
b stands for *procedures of solutions*,
c stands for *examples*.

Therefore, *Calculus 1a* means that this note is the first one on Calculus, where *a* indicates that it is a compendium.

Leif Mejlbro
9th May 2014

4

1 Complex Numbers

1.1 Introduction.

Although the main subject is real functions it would be quite convenient in the beginning to introduce the complex numbers and the complex exponential function. This is the reason for this chapter, which otherwise apparently does not seem to fit in.

The extension of the real numbers $\mathbb{R}$ to the complex numbers $\mathbb{C}$ was carried out centuries ago, because one thereby obtained that every polynomial $P_n(z)$ of degree n has precisely n complex roots, when these are counted by multiplicity.

It soon turned up, however, that this extension had many other useful applications, of which we only mention the most elementary ones in this chapter.

1.2 Definition.

Let x, $y \in \mathbb{R}$ be real numbers. We define

(1) $z = x + \mathrm{i}\, y \in \mathbb{C}, \qquad x = \operatorname{Re} z \quad \text{and} \quad y = \operatorname{Im} z,$

where $\mathrm{i} \in \sqrt{-1}$ is an adjoint element added to $\mathbb{R}$, where

$$\mathrm{i}^2 = \mathrm{i} \cdot \mathrm{i} = -1 \qquad \text{and} \qquad \mathrm{i}^2 = -1.$$

Thus, the equation of degree two,

$$z^2 + 1 = 0 \qquad \text{or equivalently} \qquad z^2 = -1,$$

has the set of solutions

$$\sqrt{-1} := \{\mathrm{i}, -\mathrm{i}\},$$

where the symbol $\sqrt{-1}$ is considered as a set and not as a single number.

The presentation (1) of a given complex number $z \in \mathbb{C}$ is uniquely determined by its real and imaginary parts.

Remark 1.1 Notice that in spite of the name *"imaginary part"*, $y = \operatorname{Im} z \in \mathbb{R}$ is always *real.* It is the uniquely determined coefficient to $\mathrm{i} \in \mathbb{C}$ in the presentation (1). $\Diamond$

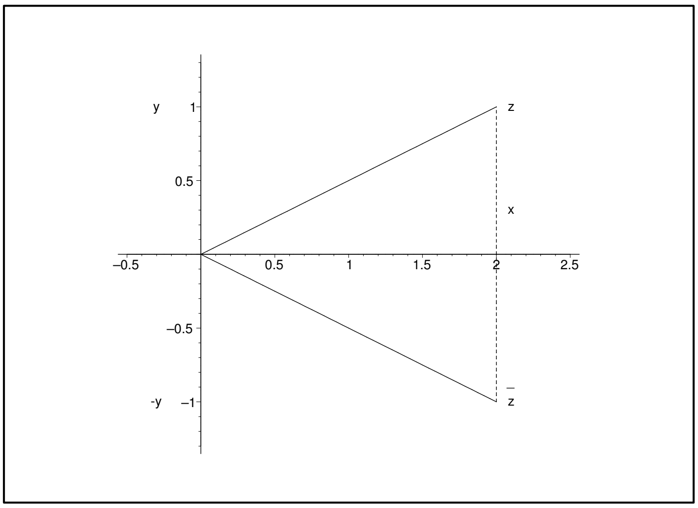

Figure 1: Rectangular description of complex conjugation.

1.3 Rectangular description in the Euclidean plane.

Let $z = x + \mathrm{i}\,y \in \mathbb{C}$, where $x,\, y \in \mathbb{R}$. Then z corresponds to the point $(x, y) \in \mathbb{R}^2$, i.e.

$$\mathbb{C} \ni z = x + \mathrm{i}\,y \sim (x, y) \in \mathbb{R}^2.$$

This correspondence is one-to-one, and we see that

$$1 \sim (1, 0) \qquad \text{and} \qquad \mathrm{i} \sim (0, 1)$$

correspond to the unit vectors on the X- axis and the Y-axis, respectively.

We introduce the *complex conjugation* of $z = x + \mathrm{i}\,y$, $x,\, y \in \mathbb{R}$, by

$$\overline{z} = x - \mathrm{i}\,y \in \mathbb{C},$$

i.e. the corresponding point $(x, y) \in \mathbb{R}^2$ is reflected in the X-axis to $\overline{z} \sim (x, -y) \in \mathbb{R}^2$.

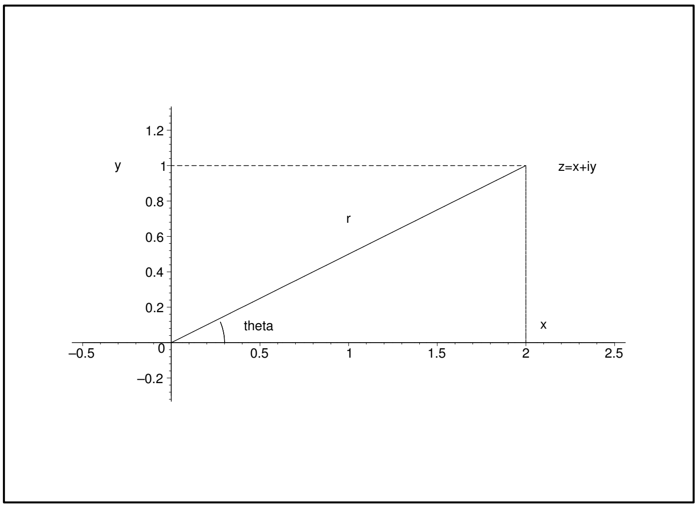

Figure 2: Polar coordinates.

1.4 Description of complex numbers in polar coordinates.

There is an *alternative* useful way of describing points in the Euclidean plane, namely by using *polar coordinates.*

Let $z = x + \mathrm{i}\,y \neq 0$ correspond to the point $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$. Then

(2) $z = x + \mathrm{i}\,y = r \cdot \cos\theta + \mathrm{i}\,r \cdot \sin\theta = r\{\cos\theta + \mathrm{i} \cdot \sin\theta\}$,

where

$$r := \sqrt{x^2 + y^2} = |z|$$

is called the *modulus* (or the *absolute value* or the *norm*) of z, and the angle θ from the positive X-axis to the line from the *pole* $(0, 0)$ to $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$, using the usual orientation is called an *argument* of z.

Hence, we have the correspondences

(3) $x = r \cdot \cos\theta, \qquad y = r \cdot \sin\theta, \qquad r > 0,$

and

(4) $r = \sqrt{x^2 + y^2} = |z|, \qquad \cos\theta = \dfrac{x}{x^2 + y^2}, \qquad \sin\theta = \dfrac{y}{x^2 + y^2},$

between rectangular coordinates $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ and polar coordinates $(r, \theta) \in \mathbb{R}_+ \times \mathbb{R}$.

Finally, we describe $0 = 0 + \mathrm{i} \cdot 0 \sim (0, 0)$ in polar coordinates by $r = 0$ and $\theta \in \mathbb{R}$ arbitrary. It is seen by (2) that any of the polar coordinates $(r, \theta) = (0, \theta)$ give the complex number $z = 0$.

The immediate problem with polar coordinates is that the *argument* is not uniquely determined. If θ_0 is an argument of a complex number $z \neq 0$, then any number $\theta_0 + 2p\pi$, where $p \in \mathbb{Z}$ is a positive or negative integer, is also an argument of the given $z \neq 0$.

This uncertainty modulo 2π with respect to the integers $\mathbb{Z}$ makes the students a little uneasy by their first encounter with polar coordinates. We shall soon see that this uncertainty actually is an advantage.

But first we have to describe the rules of operations.

1.5 Algebraic operations in rectangular coordinates.

We assume in the following that

$$z = x + \mathrm{i}\, y \in \mathbb{C}, \quad x, y \in \mathbb{R}, \quad \text{and} \quad w = u + \mathrm{i}\, v \in \mathbb{C}, \quad u, v \in \mathbb{R}.$$

Addition. Define

$$z + w = (x + \mathrm{i}\, y) + (u + \mathrm{i}\, v) := \{x + u\} + \mathrm{i}\,\{y + v\},$$

(real part plus real part, and imaginary part plus imaginary part). This corresponds to *addition of coordinates* in $\mathbb{R}^2$, and to the *rule of the parallelogram of forces* in Physics.

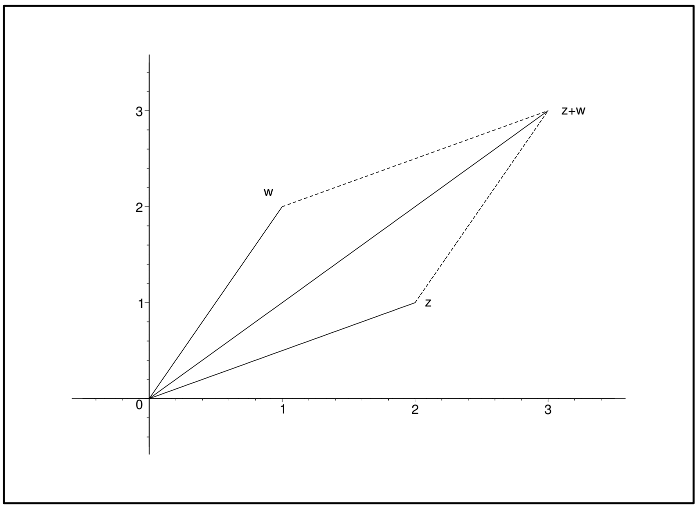

Figure 3: Addition of two complex numbers.

Subtraction. First define the *inverse* with respect to addition by

$$-z = -(x + \mathrm{i}\, y) := -x - \mathrm{i}\, y = (-x) + \mathrm{i}(-y).$$

Then *subtraction* is obtained by the composition

$$z - w := z + (-w) = (x + \mathrm{i}\, y) + (-u - \mathrm{i}\, v) = \{x - u\} + \mathrm{i}\,\{y - v\}.$$

Multiplication. Using that $\mathrm{i}^2 = -1$, *multiplication* is defined by

$$\begin{aligned} z \cdot w &= (x + \mathrm{i}\, y) \cdot (u + \mathrm{i}\, v) := xu + \mathrm{i}\,(yu + xv) + \mathrm{i}^2 yv \\ &= \{xu - yv\} + \mathrm{i}\,\{xv + yu\}. \end{aligned}$$

Hence, the *real part of the product* is "the real part times the real part *minus* the imaginary part times the imaginary part", and the *imaginary part of the product* is "the real part times the imaginary part *plus* the imaginary part times the real part".

Complex conjugation of a complex number $z = x + \mathrm{i}\, y \in \mathbb{C}$, x, $y \in \mathbb{R}$, is defined by

$$\overline{z} = x - \mathrm{i}\, y,$$

which corresponds to an orthogonal reflection of $z \sim (x, y)$ in the real axis.

Note in particular that

$$z\,\overline{z} = (x + \mathrm{i}\, y)(x - \mathrm{i}\, y) = x^2 - \mathrm{i}^2 y^2 = x^2 + y^2 := |z|^2,$$

where

$$|z| := \sqrt{x^2 + y^2}$$

is the *norm* (or the *absolute value* of the *modulus*) of $z \in \mathbb{C}$.
The norm $|z - w|$ describes the Euclidean distance in the corresponding plane $\mathbb{R}^2$ between z and w.

Obviously, $|z| \geq 0$ and

$$|z \cdot w| = |z| \cdot |w| \quad \text{and} \quad |z + w| \leq |z| + |w|.$$

Division. When $z = x + \mathrm{i}\, y \neq 0$, then also $\overline{z} \neq 0$, so we get the *inverse* of z with respect to multiplication by a small trick, namely by multiplying the numerator and the denominator by $\overline{z}$,

$$\frac{1}{z} = \frac{1}{x + \mathrm{i}y} = \frac{\overline{z}}{z \cdot \overline{z}} = \frac{x - \mathrm{i}\, y}{x^2 + y^2} = \frac{x}{x^2 + y^2} - \mathrm{i}\, \frac{y}{x^2 + y^2}.$$

In general we define *division* by some $z \neq 0$ by either of the following two methods

$$\frac{w}{z} = \begin{cases} w \cdot \dfrac{1}{z} = (u + \mathrm{i}\, v) \cdot \left\{ \dfrac{x}{x^2 + y^2} - \mathrm{i}\, \dfrac{y}{x^2 + y^2} \right\}, \\[2ex] \dfrac{w \cdot \overline{z}}{z \cdot \overline{z}} = \dfrac{(u + \mathrm{i}\, v) \cdot (x - \mathrm{i}\, y)}{x^2 + y^2}, \end{cases} \qquad z \sim (x, y) \neq (0, 0).$$

The result is of course the same, no matter which method is used, but sometimes one of them is less cumbersome to perform than the other one.

Roots of complex numbers. It is usually not possible to calculate the n-roots of a complex number in rectangular coordinates. The only possible exception is the square root where we in some cases may get even quite reasonable results.

Let $c = a + \mathrm{i}\, b$ be a complex number, and let $z = x + \mathrm{i}\, y$ be a square root of c, i.e. we have the equation

$$z^2 = c.$$

When we calculate the left hand side of this equation, we obtain

$$c = a + \mathrm{i}\, b = z^2 = (x + \mathrm{i}\, y)^2 = x^2 - y^2 + 2\mathrm{i}\, xy,$$

hence by separating the real and the imaginary parts,

$$\begin{cases} a = x^2 - y^2, \\[2ex] b = 2xy, \end{cases} \qquad \text{from which } \sqrt{a^2 + b^2} = x^2 + y^2.$$

Thus

$$x^2 = \frac{1}{2} \left\{ \sqrt{a^2 + b^2} + a \right\} > 0 \quad \text{and} \quad y^2 = \frac{1}{2} \left\{ \sqrt{a^2 + b^2} - a \right\} > 0.$$

When we apply the usual real square root we get

$$x = \pm \sqrt{\frac{\sqrt{a^2 + b^2} + a}{2}} \qquad \text{and} \qquad y = \pm \sqrt{\frac{\sqrt{a^2 + b^2} - a}{2}}.$$

This gives us four possibilities, of which only two are solutions, namely the two pairs (x, y), for which also $b = 2xy$ (check!).

In general these solutions are not simple (a square root inside a square root), but for suitable choices of the pair (a, b), which "accidentally" happen quite often in exercises, and at examinations, one nevertheless obtains simple rectangular expressions.

1.6 The complex exponential function.

A beneficial definition is

(5) $\exp(\mathrm{i}\,\theta) = e^{\mathrm{i}\,\theta} := \cos\theta + \mathrm{i}\cdot\sin\theta, \qquad \theta\in\mathbb{R}.$

The reason for this definition is given by the following calculation,

$$\begin{aligned}\exp(\mathrm{i}\,\theta_1)\cdot\exp(\mathrm{i}\,\theta_2) &:= \{\cos\theta_1 + \mathrm{i}\,\sin\theta_1\}\cdot\{\cos\theta_2+\mathrm{i}\cdot\sin\theta_2\}\\ &= \cos\theta_1\cdot\cos\theta_2 - \sin\theta_1\cdot\sin\theta_2 + \mathrm{i}\,\{\cos\theta_1\cdot\sin\theta_2+\sin\theta_1\cdot\cos\theta_2\}\\ &= \cos(\theta_1+\theta_2)+\mathrm{i}\,\sin(\theta_1+\theta_2) := \exp(\mathrm{i}\,\{\theta_1+\theta_2\}),\end{aligned}$$

so (5) satisfies the usual functional equation for the exponential function

$$\exp(\mathrm{i}\,\{\theta_1+\theta_2\}) = \exp(\mathrm{i}\,\theta_1)\cdot\exp(\mathrm{i}\,\theta_2), \qquad \theta_1,\,\theta_2\in\mathbb{R},$$

here extended to imaginary arguments.

A natural extension of (5) to general complex numbers $z\in\mathbb{C}$ is given by

$$\begin{aligned} e^z &= \exp z = \exp(x+\mathrm{i}\,y) := \exp(x)\cdot\exp(\mathrm{i}\,y) = e^x\cdot e^{\mathrm{i}\,y}\\ \text{(6)}\qquad &= e^x\{\cos y + \mathrm{i}\,\sin y\} = e^x\cos y + \mathrm{i}\,e^x\sin y.\end{aligned}$$

It follows immediately from (5) and (6) that

$$\exp(z+2\mathrm{i}\,p\pi) = \exp z, \qquad p\in\mathbb{Z},$$

so the complex exponential function is periodic with the imaginary period $2\mathrm{i}\,\pi$.

It follows from (5) that

$$e^{\mathrm{i}\,n\theta} = \cos n\theta + \mathrm{i}\,\sin n\theta = \left(e^{\mathrm{i}\,\theta}\right)^n = \{\cos\theta+\mathrm{i}\,\theta\}^n,$$

which is called *de Moivre's formula.* This can be used to derive various trigonometric formulæ.

Example 1.1 Choosing $n=3$ in *de Moivre's formula* we get by using the binomial formula,

$$\begin{aligned}\cos 3\theta + \mathrm{i}\,\sin 3\theta &= (\cos\theta+\mathrm{i}\,\sin\theta)^3\\ &= \cos^3\theta + 3\mathrm{i}\,\cos^2\theta\,\sin\theta - 3\cos\theta\,\sin^2\theta - \mathrm{i}\,\sin^3\theta\\ &= \left\{\cos^3\theta - 3\cos\theta\,\sin^2\theta\right\} + \mathrm{i}\left\{3\cos^2\theta\,\sin\theta-\sin^3\theta\right\}\\ &= \left\{4\cos^3\theta - 3\cos\theta\right\} + \mathrm{i}\left\{3\sin\theta - 4\sin^3\theta\right\}.\end{aligned}$$

By separating the real and the imaginary parts we finally get

$$\cos 3\theta = 4\cos^3\theta - 3\cos\theta, \qquad \sin 3\theta = 3\sin\theta - 4\sin^3\theta. \qquad \Diamond$$

Since

$$e^{\mathrm{i}\,\theta} = \cos\theta + \mathrm{i}\,\sin\theta \quad\text{and}\quad e^{-\mathrm{i}\,\theta} = \cos\theta - \mathrm{i}\,\sin\theta,$$

we get *Euler's formulæ*:

(7) $\cos\theta = \dfrac{1}{2}\left\{e^{\mathrm{i}\,\theta}+e^{-\mathrm{i}\,\theta}\right\} \quad\text{and}\quad \sin\theta = \dfrac{1}{2\mathrm{i}}\left\{e^{\mathrm{i}\,\theta}-e^{-\mathrm{i}\,\theta}\right\}.$

Notice also that

$$\left|e^{\mathrm{i}\,\theta}\right| = |\cos\theta + \mathrm{i}\,\sin\theta| = \cos^2\theta+\sin^2\theta = 1 \qquad \text{for all } \theta\in\mathbb{R},$$

Hence, $z = e^{\mathrm{i}\,\theta}$ describes a point on the unit circle, given in polar coordinates by $(1,\theta)$.

1.7 Algebraic operations in polar coordinates.

Let

(8) $z_1 = r_1 e^{\mathrm{i}\,\theta_1}$ and $z_2 = r_2 e^{\mathrm{i}\,\theta_2}$, $\qquad r_1,\, r_2 \geq 0, \quad \theta_1,\, \theta_2 \in \mathbb{R},$

be two complex numbers with polar coordinates (r_i, θ_i). Since

$$e^{\mathrm{i}\,\theta} := \cos\theta + \mathrm{i}\sin\theta \sim (\cos\theta, \sin\theta) \in \mathbb{R},$$

describes a unit vector of argument θ, we shall also call (8) a description of z_i in polar coordinates.

Addition and subtraction. These two operations are absolutely *not* in harmony with the polar coordinates, and even if it is possible to derive explicit formulæ, they are not of any practical use, so they shall not be given here. *Use rectangular coordinates instead!*

Multiplication and division are *better suited for polar coordinates* than for rectangular coordinates, since

$$z_1 \cdot z_2 = \left\{ r_1 e^{\mathrm{i}\,\theta_1} \right\} \cdot \left\{ r_2 e^{\mathrm{i}\,\theta_2} \right\} = (r_1\, r_2) \cdot e^{\mathrm{i}\{\theta_1 + \theta_2\}}$$

and

$$\frac{z_1}{z_2} = \frac{r_1}{r_2} \cdot e^{\mathrm{i}\{\theta_1 - \theta_2\}} \qquad \text{for } r_2 > 0.$$

For e.g. the product we multiply the moduli and add the arguments.

Complex conjugation is also an easy operation in polar coordinates:

$$\overline{z} = \overline{r\, e^{\mathrm{i}\,\theta}} = r \cdot \overline{e^{\mathrm{i}\,\theta}} = r \cdot e^{\overline{\mathrm{i}\,\theta}} = r \cdot e^{-\mathrm{i}\,\theta}.$$

Roots should *always* be calculated in polar coordinates. One needs only to remember that the complex exponential function is *periodic with the imaginary period* $2\mathrm{i}\,\pi$. Therefore, start *always* by writing

$$z = r\, e^{\mathrm{i}\,\theta} = r \cdot \exp(\mathrm{i}\{\theta + 2p\pi\}), \qquad r \geq 0, \quad \theta \in \mathbb{R}, \quad p \in \mathbb{Z},$$

where we use that

$$e^{2\mathrm{i}\,p\pi} = 1, \qquad p \in \mathbb{Z}.$$

We define for any $n \in \mathbb{N}$ the *point set* $\sqrt[n]{z}$ by

(9) $\sqrt[n]{z} = \left\{ \sqrt[n]{r} \cdot \exp\left(\mathrm{i} \cdot \dfrac{\theta + 2p\pi}{n} \right) \;\middle|\; p = 0,\, 1,\, \dots,\, n-1 \right\},$

where on the right hand side $\sqrt[n]{r}$ of a positive real number $r > 0$ is always defined as a positive real number. It is seen that we in general consider $\sqrt[n]{z}$ as a set of n complex numbers lying on the circle of centre 0 and radius $\sqrt[n]{r} \geq 0$ (real root) with the angle $\dfrac{2\pi}{n}$ between any two roots in succession.

The latter property can in some cases be used when one solves the binomial equation

$$z^n = c, \qquad \text{i.e.} \qquad z = \sqrt[n]{c}.$$

One starts by finding one solution and then turn it the angle $\dfrac{2\pi}{n}$ etc. in the plane $\mathbb{R}^2$, until one returns back to the first solution.

1.8 Roots in polynomials.

As mentioned in the Introduction (Section 1.1) every polynomial $P_n(z)$ of degree n has precisely n complex roots when we count them by multiplicity. Apart from the order of the factors, this means that any polynomial has a unique description in the form

(10) $P_n(z) = a\,(z-z_1)^{n_1}\,(z-z_2)^{n_2}\cdots(z-z_k)^{n_k}\,,$

where

$$n_1+n_2+\cdots+n_k=n, \qquad n_1,\,\dots,\,n_k\in\mathbb{N},$$

the *different* roots are the complex numbers $\{z_1,\,\dots,\,z_k\}$, each root z_j of multiplicity n_j, and $a\in\mathbb{C}$ is constant.

It follows immediately from (10) that $z=z_j$ is a root, if and only if $z-z_j$ is a divisor in $P_n(z)$, i.e. if and only if

$$P_{n-1}(z)=\begin{cases}\dfrac{P_n(z)}{z-z_j} & \text{for } z\neq z_n,\\[2ex] \lim_{z\to z_j}\dfrac{P_n(z)}{z-z_j}, & \text{for } z=z_j,\end{cases} \qquad \text{is also a polynomial (of degree } n-1).$$

All this may look captivating easy to perform, *but it is not in practice!* Therefore, use as a rule of thumb (with very few obvious exceptions) the following method: *Identify first as many factors of the form* $(z - z_j)^{n_j}$ as possible, in order to get as close as possible to the ideal form (10). There may of course still be a residual polynomial left after this procedure.

Notice also that the multiplication in (10) is only carried out in extremely rare cases, because one loses information by this operations. Therefore, *leave as many factors as possible!*

The first problem is whether one can find a formula for the roots of $P_n(z)$.

Such a formula exists when $n = 2$, and it is used over and over again in the applications. This is well-known from high school, when the coefficients are real. We shall now prove it, when the coefficients are complex.

Let a, b, $c \in \mathbb{C}$ be complex numbers, where $a \neq 0$, and let

$$P_2(z) = az^2 + bz + c.$$

If z is a root of $P_2(z)$, then we have

$$\begin{aligned} 0 &= P_2(z) = az^2 + bz + c = a\left\{z^2 + \frac{b}{a}\,z + \frac{c}{a}\right\} \\ &= a\left\{z^2 + 2\cdot\frac{b}{2a}\cdot z + \left(\frac{b}{2a}\right)^2 - \left(\frac{b}{2a}\right)^2 + \frac{c}{a}\right\} \\ &= a\left\{\left(z + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a^2}\right\}, \end{aligned}$$

from which

$$z + \frac{b}{2a} = \pm\frac{1}{2a}\sqrt{b^2 - 4ac},$$

and we have proved that the well-known formula also holds for complex coefficients,

$$(11)\quad z = \frac{1}{2a}\left\{-b \pm \sqrt{b^2 - 4ac}\right\}.$$

There exist solution formulæ (*Cardano's formula* and *Vieti's formula*) for $n = 3$, but these are very complicated for practical use, so they are not given here.

There is also an exact solution procedure when $n = 4$, but it is relying on the solution for $n = 3$, so this solution formula is not given either.

And then came the great shock for the mathematicians of that time, when it was proved nearly two centuries ago that *there exists no general solution formula for the roots of* $P_n(z)$ *when* $n \geq 5$. Hence, we can no longer be sure to find the exact roots of a polynomial of degree ≥ 5. This is another reason for not giving the exact formulæ for $n = 3$ and $n = 4$, even though they exist.

Polynomials and their roots are incredibly important in the applications. Therefore, in the remainder part of this chapter we give some methods by which we may be able to find some of the roots. Notice that some of these methods may no longer be found in the elementary books of Calculus.

1) If $z_0 \in \mathbb{C}$ is a given root of $P_n(z)$, i.e. $P_n(z_0) = 0$, we reduce the investigation to a polynomial of degree $n-1$ by calculating

$$P_{n-1}(z) = \frac{P_n(z)}{z - z_0}, \quad \text{so} \quad P_n(z) = (z - z_0)\, P_{n-1}(z),$$

and trivially extend this definition to $z = z_0$.

2) If $n = 2$ we of course use the solution formula (11).

3) If $P_n(z)$ has *real* coefficients, and $z_0 = x_0 + \mathrm{i}\, y_0 \in \mathbb{C}$, $y_0 \neq 0$, is a *complex root*, then the conjugate $\overline{z_0} = x_0 - \mathrm{i}\, y_0$ is also a complex root, and

$$(z - z_0)(z - \overline{z_0}) = (z - x_0)^2 + y_0^2 = z^2 - 2x_0 z + x_0^2 + y_0^2$$

is a divisor in $P_n(z)$ of real coefficients, so we can continue the investigation on the polynomial

$$P_{n-1}(z) = \frac{P_n(z)}{(z - x_0)^2 + y_0^2}$$

of degree $n-2$, and of real coefficients.

Note that if $z_0 \in \mathbb{C} \setminus \mathbb{R}$ is a root of multiplicity n_0, and $P_n(z)$ has *real* coefficients, then $\overline{z_0} \in \mathbb{C} \setminus \mathbb{R}$ is also a root of multiplicity n_0.

4) If the coefficients of the polynomial

(12) $P_n(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0, \qquad a_0,\, a_1,\, \dots,\, a_n \in \mathbb{Z},$

are all *integers*, then a rational root $z_0 \in \mathbb{Q}$, if any, must have the structure $\pm\dfrac{p}{q} \in \mathbb{Q}$, where p and $q \in \mathbb{N}$, and p is a divisor in a_0, and q is a divisor in $a_n \neq 0$.
This gives a finite number of possible rational roots, which can easily be checked.

The same procedure can be used when the coefficients all are rational numbers. First multiply the polynomial by the smallest positive integer, for which the multiplied polynomial has integers as coefficients and then apply the method described above.

5) If $P_n(z)$ has a root z_0 of multiplicity $n_0 \geq 2$, then if follows by a differentiation of (10) that z_0 is a root in the derived polynomial $P_n'(z)$ of multiplicity $n_0 - 1 \geq 1$.

This can be applied in the following way: Since $P_n(z)$ of degree n and $P_n'(z)$ of degree $n-1$ are known, we get by division

$$P_n(z) = Q_1(z) P_n'(z) + R_1(z),$$

where $R_1(z)$ is a residual polynomial of degree $< n-1$. Since $(z - z_0)^{n_0 - 1}$ is a divisor in both $P_n(z)$ and $P_n'(z)$, it must also be a divisor in the residual polynomial $R_1(z)$.

Proceed by performing the division

$$P_n'(z) = Q_2(z) R_1(z) + R_2(z),$$

where $(z - z_0)^{n_0 - 1}$ also is a divisor in the residual polynomial $R_2(z)$ of degree $< n-2$, etc.

After a final number of steps we obtain a residual polynomial $R(z) \neq 0$, such that the next residual polynomial is the zero polynomial.

a) If $R(z)$ is a constant, then all roots of $P_n(z)$ are simple.
b) If z_0 is a root of $R(z)$ of multiplicity $p \geq 1$, then z_0 is a root of $P_n(z)$ of multiplicity $p+1$.
c) If $P_n(z)$ is given by (10), then $R(z)$ is given by

$$R(z) = A\,(z-z_1)^{n_1-1}\,(z-z_2)^{n_2-1}\cdots(z-z_k)^{n_k-1},$$

where $A \in \mathbb{C}$ is some constant. Hence, the roots of $R(z)$ are all the multiple roots of $P_n(z)$.

If $P_n(z)$ does not have multiple roots, we get that $R(z)$ is a constant, and this method cannot give us further information.

On the other hand, if $R(z)$ again is of a high degree, such that one cannot immediately find its roots, we may apply the same method on $R(z)$ and $R'(z)$, thereby finding the roots in $P_n(z)$ of multiplicity ≥ 2, etc..

Notice that a clever handling of this method will give us the polynomials, the simple roots of which are precisely the roots of multiplicity 1, 2, etc.. We leave it to the reader to derive this result from the above.

6) When all other methods fail, and $P_n(z)$ has *real* coefficients, then the following method, called the *Newton-Raphson iteration method*, may be used to give approximations of the *simple, real roots* with arbitrary small error. Notice that if the root had multiplicity ≥ 2, then we would probably have found it under the method described in 5).

When we use the present method we need either a computer or at least an advanced pocket calculator. Start with an analysis on the computer of the graph of $P_n(x)$, $x \in \mathbb{R}$, in order roughly to estimate where some simple root $x_0 \in \mathbb{R}$ is located.

a) Based on this rough estimate of the simple root $x_0 \in \mathbb{R}$ (where $P_n'(x_0) \neq 0$) choose an element $x_1 \in \mathbb{R}$ close to the unknown $x_0 \in \mathbb{R}$.
b) Define the next element by

$$x_2 := x_1 - \frac{P_n(x_1)}{P_n'(x_1)}.$$

c) Define by induction

$$x_{m+1} := x_m - \frac{P_n(x_m)}{P_n'(x_m)}, \qquad m \in \mathbb{N}.$$

If x_1 is chosen sufficiently close to the root x_0, then it can be proved that

$$\lim_{m\to\infty} x_m = x_0.$$

7) On very rare occasions we get the information that $P_n(z)$ has a real root $x_0 \in \mathbb{R}$, and that not all coefficients of $P_n(z)$ are real. In this case $x - x_0$ must be a divisor in both Re $P_n(x)$ and Im $P_n(x)$, $x \in \mathbb{R}$, so we can use the same procedure on the pair $(\operatorname{Re} P_n(x), \operatorname{Im} P_n(x))$ as we did on the pair $(P_n(z), P_n'(z))$ in the method described in 5). Thus, if Re $P_n(x)$ has at least the same degree as Im $P_n(z)$, then

$$\operatorname{Re} P_n(x) = Q_1(x) \cdot \operatorname{Im} P_n(x) + R_1(x),$$

where $x - x_0$ is a divisor in $R_1(x)$, which is of smaller degree than Im $P_n(x)$, etc..

8) A similar method can in some cases be used to find complex conjugated roots in a polynomial of real coefficients. An example is the equation

$$z^5 - 1 = 0,$$

where one puts $z = x + \mathrm{i}\, y$ and then for $y \neq 0$ splits the resulting computed equation into its real part and its imaginary part.

Warning. Even in the simple example mentioned above the necessary calculations are really tough, so this method cannot in general be recommended.

2 The Elementary Functions

2.1 Introduction.

The elementary functions and their properties form the simplest and yet the most important building stones in Calculus. They are in general extremely boring to be studied isolated, but they are on the other hand also extremely important in the practical applications. They *must* therefore be mastered by the student before he or she can proceed to more advanced problems.

We shall here in general describe the functions by the following scheme, whenever it makes sense:

- Definition
- Graph
- Derivative
- Rules of algebra etc.

For more detailed formulæ etc. the reader is referred to the appendix.

Since this is not meant to be an ordinary textbook we shall start untraditionally by describing the concept of an *inverse function*.

2.2 Inverse functions.

We shall only consider the simplest case where there is given a *continuous, strictly monotonous function* $f : I \to J$ of an interval I onto another interval J.

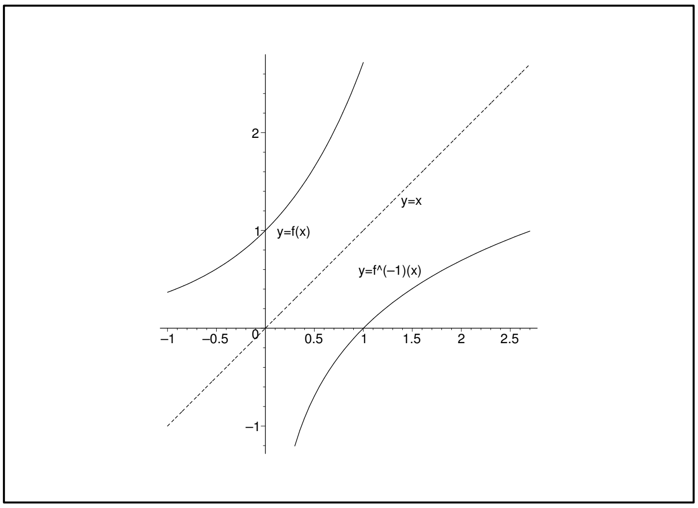

Figure 4: The graph of the inverse function is obtained by reflecting the graph of the function in the line $y = x$.

Since $f(I) = J$ and f is strictly monotonous, we can to every $y \in J$ find one and only one $x \in I$, such that

$$f(x) = y.$$

In this way we have defined a function $\varphi : J \to I$, or $f^{-1} : J \to I$, by

(13) $x = \varphi(y)$ or $x = f^{-1}(y)$, when $f(x) = y$.

The function $\varphi = f^{-1}$ is called the *inverse function* of f.

When the inverse function of f exists, the graph of f^{-1} is obtained by "interchanging the X-axis and the Y-axis", which can also be described by reflecting the graph of f perpendicularly in the line $y = x$, cf. the figure.

It follows from this geometrical property that if $y = f(x)$ is strictly monotonous and differentiable with $f'(x) \neq 0$ for all $x \in I$, then the inverse function $x = \varphi(t)$ is also differentiable, and the derivative is given by

(14) $\varphi'(y) = \dfrac{1}{f'(\varphi(y))}$.

This follows from the fact that if a tangent has the slope $\alpha \neq 0$ in the XY-plane, then the same tangent has the slope $\dfrac{1}{\alpha}$ in the YX-plane, where the axes are interchanged.

Alternatively we have

$$y = f(x) = f(\varphi(y)),$$

so by differentiating the composite function with respect to y we get

$$1 = f'(\varphi(y)) \cdot \varphi'(y),$$

and (14) follows by a rearrangement.

Mnemonic rule. The important formula (14) is remembered by the formally incorrect equation

$$\frac{dx}{dy} \cdot \frac{dy}{dx} = 1, \qquad \text{i.e.} \quad \frac{dx}{dy} = \frac{1}{\frac{dy}{dx}},$$

hence

$$\frac{dx}{dy} = \varphi'(y) \qquad \text{and} \qquad \frac{1}{\frac{dy}{dx}} = \frac{1}{f'(x)} = \frac{1}{f'(\varphi(y))},$$

and (14) follows.

Notice that if one can find two different points x_1 and $x_2 \in I$, such that

$$f(x_1) = f(x_2) \qquad (= y \in J),$$

then f does not have an inverse function.

However, in some cases one can find a subinterval $I_1 \subset I$, such that $f : I_1 \to J$ is *one-to-one*, thus in this case we can find a "local" inverse function $\varphi_1 : J \to I_1$. This is typically the case when we shall define reasonable inverse functions of the trigonometric functions.

2.3 Logarithms and exponentials.

The simplest family of functions consists of the polynomials, and one would usually start with them in textbooks. Here it will give a better presentation if we instead start with the (natural) *logarithm* and the *exponential function*.

Definition 2.1 *The (natural)* logarithm $\ln : \mathbb{R}_+ \to \mathbb{R}$ *is defined by*

(15) $\ln x = \displaystyle\int_1^x \frac{1}{t}\,dt, \qquad x > 0.$

If follows from (15) that

(16) $\dfrac{d}{dx}\ln x = \dfrac{1}{x} > 0, \qquad x > 0,$

so $\ln x$ is continuous and strictly increasing.

It is well-known that $\ln(\mathbb{R}_+) = \mathbb{R}$, so the inverse function "$\ln^{-1} : \mathbb{R} \to \mathbb{R}_+$" exists. It is denoted by $\exp : \mathbb{R} \to \mathbb{R}_+$, and it is called the *exponential function*. Hence,

$$y = \ln x, \qquad \text{if and only if} \qquad x = \exp y, \qquad x \in \mathbb{R}_+ \text{ and } y \in \mathbb{R}.$$

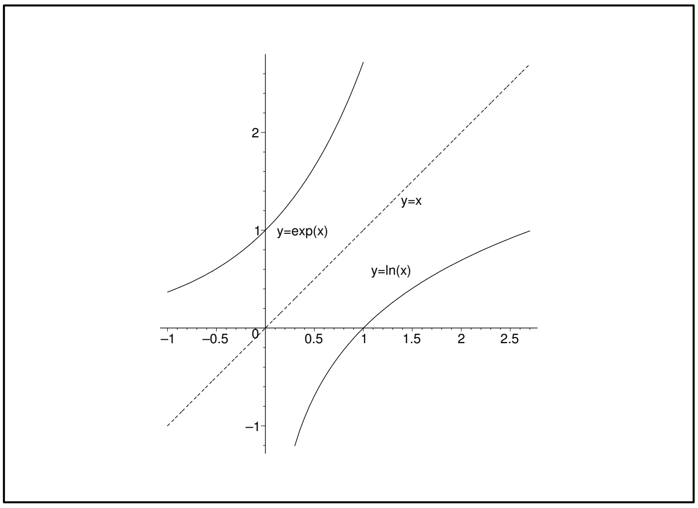

Figure 5: The graphs of $y = \ln x$ and its inverse $y = \exp x$.

It follows from (14) that the derivative is given by

$$\frac{d}{dy}\exp y = \frac{1}{1/x} = x = \exp y,$$

hence by replacing y by x,

(17) $\dfrac{d}{dx}\exp x = \exp x,$

so $\exp x$ is invariant with respect to differentiation.

Algebraic rules; functional equations. These should be well-known,

(18) $\ln(a \cdot b) = \ln a + \ln b, \qquad \ln \frac{a}{b} = \ln a - \ln b, \qquad a,\, b \in \mathbb{R}_+,$

and

(19) $\exp(c+d) = \exp(c) \cdot \exp(d), \qquad \exp(c-d) = \frac{\exp c}{\exp d}, \qquad c,\, d \in \mathbb{R}.$

Formula (18) shows that ln carries a product (quotient) of positive numbers into a sum (difference) of logarithms.

Formula (19) shows that exp carries a sum (difference) of real numbers into a product (quotient) of exponentials.

Extensions. The functions introduced above are extended in the following ways:

Logarithmic functions with base $g > 0$, $g \neq 1$. These are defined by

$$y = \log_g x := \frac{\ln x}{\ln g}, \qquad x \in \mathbb{R}_+,$$

where the derivative is

$$\frac{d}{dx}\,\log_g x = \frac{1}{x \ln g}, \qquad x \in \mathbb{R}_+.$$

When $g = 10$, the function is also called the *common logarithm* and denoted by

$$y = \log_{10} x := \log x, \qquad x \in \mathbb{R}_+.$$

The common logarithm was before the introduction of e.g. pocket calculators the most used logarithmic function, because generations of mathematicians had worked out tables of this function. Today only the *natural logarithm* is of importance. It corresponds of course to the base

$$g = e := \exp(1).$$

Exponentials with base $a > 0$. These are more important. They are defined by

$$(20)\quad y = a^x := \exp(x\,\ln a), \qquad x \in \mathbb{R},$$

with the derivative

$$\frac{da^x}{dx} = \ln a \cdot a^x.$$

Choosing $a = e := \exp(1)$, and using that ln is the inverse function, we see that $\ln e = 1$, Hence, we get in particular by (20) that

$$y = e^x := \exp(x), \qquad x \in \mathbb{R}.$$

Notice that if $a = 1$, then of course

$$1^x := 1, \qquad x \in \mathbb{R},$$

with no inverse function.

Functional equations. Whenever $a > 0$,

$$a^{x+y} = a^x \cdot a^y, \qquad a^{x-y} = \frac{a^x}{a^y}, \qquad x,\, y \in \mathbb{R}.$$

It is seen that the exponentials carry a sum (difference) into a product (quotient).

2.4 Power functions.

Let $\alpha \in \mathbb{R}$ be a constant. We define the *power function* $y = x^\alpha$ of exponent $\alpha \in \mathbb{R}$ by

$$(21)\quad y = x^\alpha := \exp(\alpha \cdot \ln x), \qquad x \in \mathbb{R}_+,$$

cf. (20)

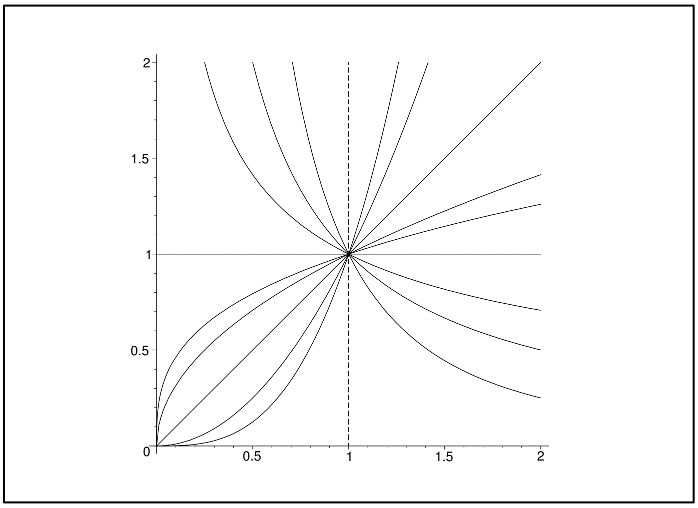

Figure 6: The graphs of the power functions $y = x^{\alpha}$, where $\alpha = -2,\ -1,\ -\frac{1}{2},\ 0,\ \frac{1}{3},\ \frac{1}{2}$, 1,2,3.

We note that when $\alpha \neq 0$, then the *inverse* of the power function $y = x^{\alpha}$ is given by

$$x = y^{1/\alpha} := \exp\left(\frac{1}{\alpha}\ln y\right),$$

which is easily checked by means of (21).

Possible extensions of the definition. It is possible for certain values of $\alpha \in \mathbb{R}$ to extend the domain of the corresponding power function.

1) If $\alpha = n \in \mathbb{N}_0$ is a nonnegative integer, the power function x^n can be considered as a product of n factors, all equal to x, i.e.

   (22) $x^n = \underbrace{x \cdots x}_{n \text{ factors}}, \qquad x^0 := 1 \quad \text{for } n = 0.$

   These expressions make sense for every $x \in \mathbb{R}$, so we can extend their domain to $\mathbb{R}$. Each function x^n is called a *monomial*, and a finite sum of monomials, e.g.

   $$P_n(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \qquad x \in \mathbb{R},$$

   where $a_0,\ a_1,\ \dots,\ a_n \in \mathbb{R}$ (or $\in \mathbb{C}$) are real (or complex) constants, and $a_n \neq 0$, is called a *polynomial of degree* $n \in \mathbb{N}$.

   Remark 2.1 The set of all polynomials form a very important class of functions. It is notable that even this fairly simple class of functions may cause serious difficulties in the computations. Cf. e.g. the various methods described in Section 1.8, which may be used to find the complex roots, if possible. $\Diamond$

2) When $\alpha = n \in \mathbb{N}$ is a positive integer, we use (22) to define x^{-n} (with a negative integer as exponent) by

   (23) $x^{-n} := \dfrac{1}{x^n}, \qquad x \in \mathbb{R} \setminus \{0\}.$

We must obviously exclude $x = 0$ from the extension, because we are never allowed to divide by 0.

3) If $\alpha \in \mathbb{R}_+$, then $x^\alpha = \exp(\alpha \cdot \ln x) \to 0$ for $x \to 0+$. In this case we define the extension by

$$x^\alpha := \begin{cases} \exp(\alpha \cdot \ln x) & \text{for } x > 0, \\ 0 & \text{for } x = 0. \end{cases}$$

4) If $\alpha = \dfrac{p}{q} \in \mathbb{Q}$ is a rational number with an odd denominator $q \in \mathbb{N}$, then we can also calculate the q-root of a negative number. We therefore define the extension by

$$x^{p/q} := \begin{cases} \exp\left(\dfrac{p}{q} \cdot \ln x\right) & \text{for } x > 0, \\ -\exp\left(\dfrac{p}{q} \cdot \ln |x|\right) & \text{for } x < 0. \end{cases}$$

If furthermore $\alpha = \dfrac{p}{q} > 0$, we add $0^{p/q} := 0$ to the extension.

It is seen that in certain cases one may use more than one of the extensions mentioned above. Therefore, in general an extension of a power function may not be as easy as one expects. The best advice is always to use common sense for any given $\alpha \in \mathbb{R}$.

Derivative. The derivative of the power function $y = x^\alpha$ is given by

(24) $$\frac{dy}{dx} = \alpha \cdot x^{\alpha-1}$$

in each of the open subintervals in which x^α is defined, including possible extensions. Note that (24) is always valid for $x > 0$, but its domain of validity may be bigger.

Functional equations. Many years of experience in teaching have shown me that even the simplest rules known from high school may cause troubles in practice.

Whenever $a, b \in \mathbb{R}_+$ and $r, s \in \mathbb{R}$ we have

(25) $$\begin{cases} a^{rs} = a^r \cdot a^s, & (a^r)^s = a^{r\cdot s}, \\ (ab)^r = a^r \cdot b^r, & \left(\dfrac{a}{b}\right)^r = \dfrac{a^r}{b^r}, \\ a^{-r} = \dfrac{1}{a^r}, & \sqrt[r]{a} = a^{1/r}, \\ \ln(a^r) = r \cdot \ln a. & \end{cases}$$

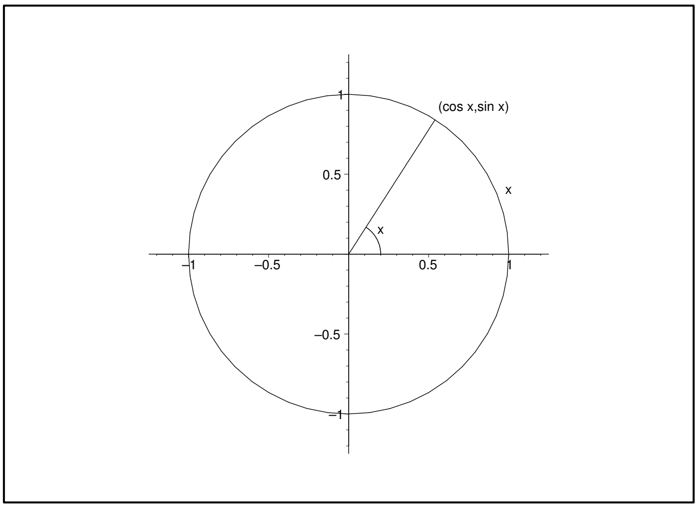

Figure 7: The geometrical meaning of $(\cos x, \sin x)$.

2.5 Trigonometric functions.

(*Trigonometry*, from Greek, "measure of triangles".) The elementary trigonometric functions are

$\cos x$, $\sin x$, $\tan x$ and $\cot z$.

The geometrical interpretation of the first two functions is that $(\cos x, \sin x)$ are the coordinates of the unit circle, where the arc length x on the unit circle is calculated from $(1, 0)$ with positive orientation (i.e. the direction from the positive X-axis to the positive Y-axis).

An equivalent description is that x denotes the angle measured in radians between the vectors $(1, 0)$ and $(\cos x, \sin x)$. This very simple geometrical interpretation is often neglected by the students, probably because it is "too obvious".

Since $(\cos x, \sin x)$ is a point on the unit circle, its Euclidean distance from the centre $(0, 0)$ is always 1. Hence we have shown

The basic trigonometric relation

(26) $\cos^2 x + \sin^2 x = 1 \qquad$ for all $x \in \mathbb{R}$.

It is obvious that the trigonometric functions $\cos x$ and $\sin x$ are particular suited for describing circular motions, and historically they were precisely defined to describe such motions. It turned up later that the trigonometric functions also were interesting for their own sake. We therefore start by drawing their well-known graphs, from which it is seen that they are periodic functions of period 2π.

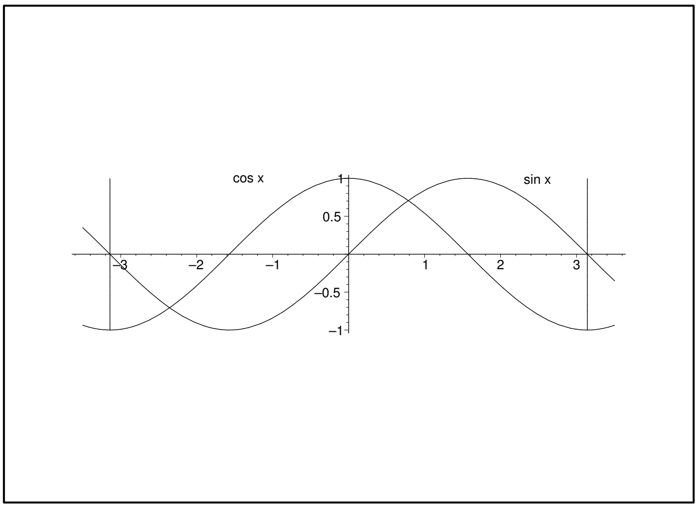

Figure 8: The graphs of $\cos x$ and $\sin x$ with an interval of periodicity $[-\pi, \pi]$.

We define two other elementary trigonometric functions, $\tan x$ and $\cot x$, by

(27) $\tan x = \dfrac{\sin x}{\cos x} \qquad$ for $x \neq \dfrac{\pi}{2} + p\pi, \quad p \in \mathbb{Z},$

(28) $\cot x = \dfrac{\cos x}{\sin x} \qquad$ for $x \neq p\pi, \qquad p \in \mathbb{Z}.$

They are periodic functions of period π.

Notice in particular that

(29) $\tan x \cdot \cot x = 1 \quad$ or $\quad \cot x = \dfrac{1}{\tan x} \qquad$ for $x \neq p \cdot \dfrac{\pi}{2}, \quad p \in \mathbb{Z},$

so it would be sufficient just to use one of them. But since they both are convenient in the applications, we shall keep both here.

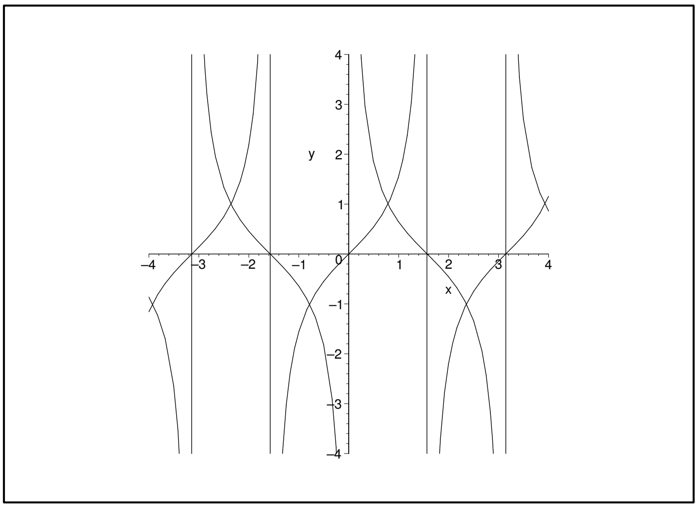

Figure 9: The graphs of $\tan x$ and $\cot x$.

The derivatives. These are

$$(30)\quad \begin{cases} \dfrac{d}{dx}\cos x = -\sin x, \quad x \in \mathbb{R}, & \dfrac{d}{dx}\tan x = \dfrac{1}{\cos^2 x} = 1 + \tan^2 x, \quad x \neq \dfrac{\pi}{2} + p\pi, \\[2ex] \dfrac{d}{dx}\sin x = \cos x, \quad x \in \mathbb{R}, & \dfrac{d}{dx}\cot z = -\dfrac{1}{\sin^2 x} = -\left(1 + \cot^2 x\right), \quad x \neq p\pi. \end{cases}$$

Functional equations. Since mathematicians have worked with trigonometric functions for centuries there exist many functional relations. Here, only the simplest addition formulæ are mentioned,

$$(31)\quad \begin{cases} \cos(x+y) = \cos x \cdot \cos y - \sin x \cdot \sin y, \\[1ex] \cos(x-y) = \cos x \cdot \cos y + \sin x \cdot \sin y, \\[1ex] \sin(x+y) = \sin x \cdot \cos y + \cos x \cdot \sin y, \\[1ex] \sin(x-y) = \sin x \cdot \cos y - \cos x \cdot \sin y. \end{cases}$$

Other functional relations can be found in the appendix.

Connection with the complex exponential function. Some of the many functional relations can easily be derived by applying the *complex exponential function*, introduced in Section 1.6 by

$$e^{\mathrm{i}\,x} = \cos x + \mathrm{i} \cdot \sin x.$$

One example is the calculation

$$\begin{aligned} e^{\mathrm{i}(x+y)} &= \cos(x+y) + \mathrm{i} \cdot \sin(x+y) = e^{\mathrm{i}\,x} \cdot e^{\mathrm{i}\,y} \\ &= \{\cos x + \mathrm{i}\,\sin x\} \cdot \{\cos y + \mathrm{i}\,\sin y\} \\ &= \cos x \cdot \cos y - \sin x \cdot \sin y + \mathrm{i} \cdot \{\sin x \cdot \cos y + \cos x \cdot \sin y\}. \end{aligned}$$

By a splitting into the real and imaginary parts we obtain again the addition formulæ,

$$\begin{aligned}\cos(x+y) &= \cos x\cdot\cos y - \sin x\cdot\sin y,\\ \sin(x+y) &= \sin x\cdot\cos y + \cos x\cdot\sin y.\end{aligned}$$

Remark 2.2 Historically it was the other way round. One *knew* the trigonometric additional formulæ, from which the complex exponential function was *defined* as also was done here previously, and its usual properties were proved. Once this has been done, it is easier to remember the functional equation for the exponential function than all these trigonometric formulæ, etc.. $\Diamond$

2.6 Hyperbolic functions.

These are denoted

$$\cosh x, \qquad \sinh x, \qquad \tanh x \quad \text{and} \quad \coth x.$$

They are analogous to the trigonometric functions and they share many of their properties, but here we can in their definitions refer directly to the real exponential function. We define

$$(32)\quad \cosh x := \frac{1}{2}\left\{e^x + e^{-x}\right\}, \qquad \sinh x = \frac{1}{2}\left\{e^x - e^{-x}\right\},$$

which are analogous to Euler's formulæ for the trigonometric functions,

(33) $\cos x = \frac{1}{2}\left\{e^{\mathrm{i}\,x} + e^{-\mathrm{i}\,x}\right\}, \qquad \sin x = \frac{1}{2\mathrm{i}}\left\{e^{\mathrm{i}\,x} - e^{-\mathrm{i}\,x}\right\}.$

It follows from (32) and (33) that

$$\cosh(\mathrm{i}\,x) = \frac{1}{2}\left\{e^{\mathrm{i}\,x} + e^{-\mathrm{i}\,x}\right\} = \cos x,$$

and

$$\sinh(\mathrm{i}\,z) = \frac{1}{2}\left\{e^{\mathrm{i}\,x} - e^{-\mathrm{i}\,x}\right\} = \mathrm{i}\,\sin x,$$

which show the close relationship between the trigonometric functions and the hyperbolic functions.

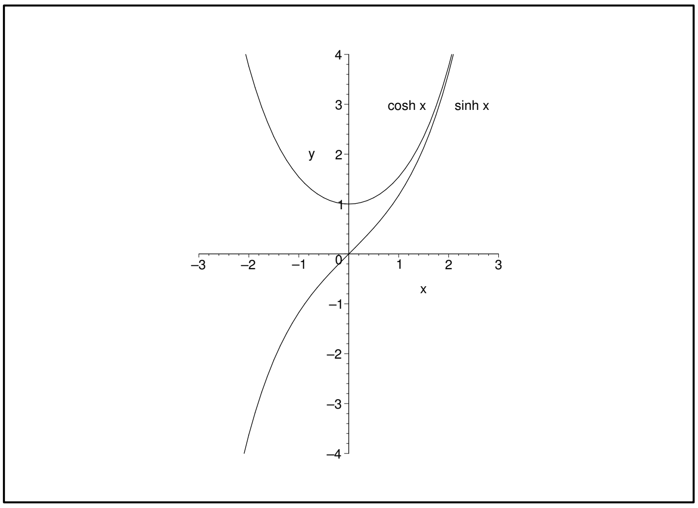

Figure 10: The graphs of $\cosh t$ and $\sinh t$.

By using (32) we define

(34) $\tanh x := \dfrac{\sinh x}{\cosh x} = \dfrac{e^{2x}-1}{e^{2x}+1}, \qquad x \in \mathbb{R},$

and

(35) $\coth x = \dfrac{\cosh x}{\sinh x} = \dfrac{e^{2x}+1}{e^{2x}-1}, \qquad x \in \mathbb{R}\setminus\{0\}.$

Since it follows from (32) that

(36) $\cosh x + \sinh x = e^{x}$ and $\cosh x - \sinh x = e^{-x}$,

it is obvious that we have the *basic hyperbolic relation*

(37) $\cosh^2 x - \sinh^2 x = 1.$

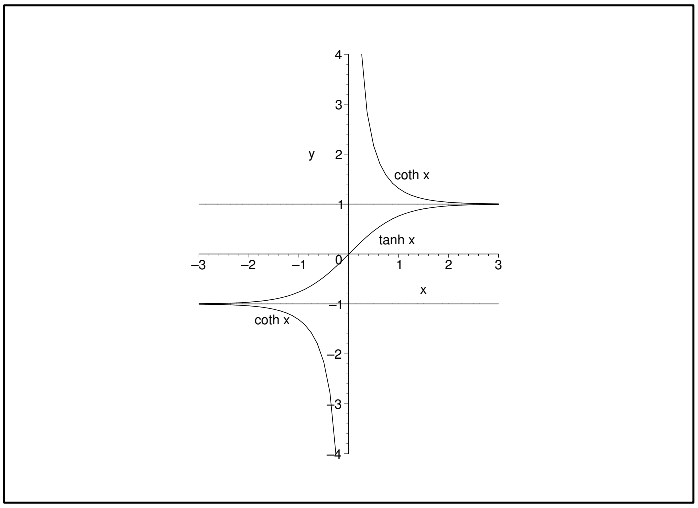

Figure 11: The graphs of $\tanh t$ and $\coth t$ and their asymptotes.

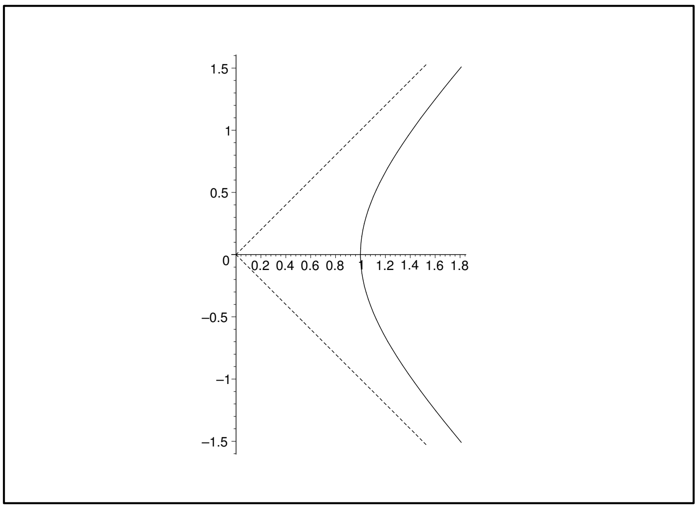

Figure 12: the graph of the curve $(\cosh t, \sinh t)$, $t \in \mathbb{R}$, is a part of the unit hyperbola $x^2 - y^2 = 1$.

A geometrical interpretation of (37) says that the point $(\cosh t, \sinh t)$, $t \in \mathbb{R}$, lies on a branch of the unit hyperbola

$$x^2 - y^2 = 1$$

in the right hand side of the plane.

Derivatives. Since the hyperbolic functions can be expressed by means of the real exponential function, they are easy to differentiate, even when one has forgotten their derivatives, which are given here,

$$(38)\quad \begin{cases} \dfrac{d}{dx}\cosh x = \sinh x, & \dfrac{d}{dx}\tanh x = \dfrac{1}{\cosh^2 x} = 1 - \tanh^2 x, \\[2ex] \dfrac{d}{dx}\sinh x = \cosh z, & \dfrac{d}{dx}\coth x = -\dfrac{1}{\sinh^2 x} = 1 - \coth^2 x, \quad x \neq 0. \end{cases}$$

Functional relations. These can e.g. be derived from the definitions. They are of course analogous to the functional relations for the trigonometric functions,

$$(39)\quad \begin{cases} \cosh(x+y) = \cosh x \cdot \cosh y + \sinh x \cdot \sinh y, \\[1ex] \cosh(x-y) = \cosh x \cdot \cosh y - \sinh x \cdot \sinh y, \\[1ex] \sinh(x+y) = \sinh x \cdot \cosh y + \cosh x \cdot \sinh y, \\[1ex] \sinh(x-y) = \sinh x \cdot \cosh y - \cosh x \cdot \sinh y. \end{cases}$$

2.7 Area functions.

These are the inverse functions of the hyperbolic functions. However, since $\cosh x = \cosh(-x)$, it follows that $\cosh x$ is not monotonous in its full domain, which also is seen from the figure. We are therefore forced to restrict ourselves, e.g. to

$$\cosh : [0, +\infty[\to [1, +\infty[,$$

where the function is one-to-one, so its inverse function exists.

There is no such problem for the inverse functions of $\sinh x$, $\tanh x$ and $\coth s$.

The inverse functions to the hyperbolic functions are here denoted by

$$\mathrm{Arcosh}\, x, \quad \mathrm{Arsinh}\, x, \quad \mathrm{Artanh}\, x, \quad \mathrm{Arcoth}\, x,$$

We mention that other notations can also be found, like e.g.

$$\begin{array}{llll} \mathrm{arcosh}\, x, & \mathrm{arsinh}\, x & \mathrm{artanh}\, x & \mathrm{arcoth}\, x \\ \cosh^{-1} x, & \sinh^{-1} x, & \tanh^{-1} x, & \coth^{-1} x, \\ \mathrm{arccosh}\, x, & \mathrm{arcsinh}\, x, & \mathrm{arctanh}\, x, & \mathrm{arccoth}\, x. \end{array}$$

It is recommended to avoid the alternative notations of the latter two lines. For instance, $\cosh^{-1} x$ may be wrongly interpreted as $1/\cosh x$, and arccosh x is a hybrid of "arcus" (meaning "arc", where it should be "area") and a hyperbolic function, represented by the extra "h".

The reason for in textbooks to handle the inverse functions of the hyperbolic functions before the inverse functions of the trigonometric functions is of course that here it is possible to find an alternative explicit expression containing the logarithm and the square root. We shall show the method by finding the inverse function of $y = \cosh x$, in which case we should also make a restriction of the domain. The derivation of the inverse functions of the other hyperbolic functions is left to the reader.

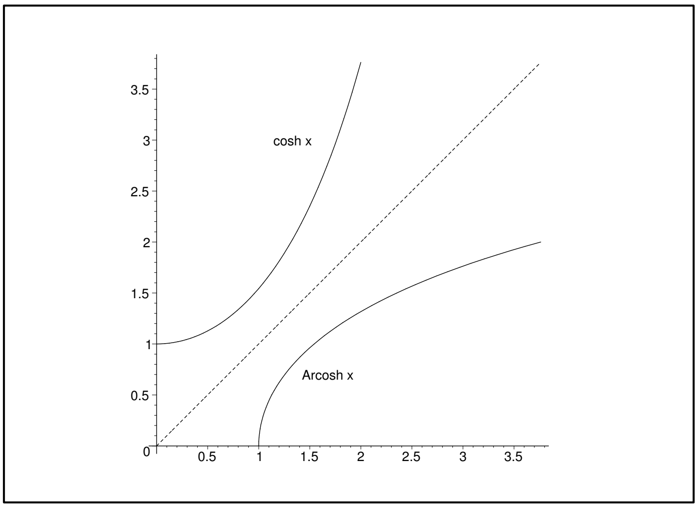

Figure 13: The graphs of $y = \cosh x$, $x \geq 0$, and $y = \operatorname{Arcosh} x$, $x \geq 1$.

The inverse function of $y = \cosh x$. When we consider the graph we conclude that the most natural restriction is given by

$$\cosh : [0, +\infty[\to [1, +\infty[.$$

The task is now after interchanging the letters to solve the equation

$$x = \cosh y, \qquad x \geq 1 \text{ and } y \geq 0.$$

Since

$$x = \cosh y = \frac{1}{2}\left\{e^y + e^{-y}\right\} = \frac{1}{2e^y}\left\{(e^y)^2 + 1\right\},$$

we get by a rearrangement,

$$(e^y)^2 - 2x \cdot e^y + 1 = 0, \qquad x \geq 1 \text{ and } y \geq 0.$$

When $x^2 - 1 \geq 0$ is added to this equation, we get

$$x^2 - 1 = (e^y)^2 - 2x \cdot e^y + x^2 = \{e^y - x\}^2, \qquad x \geq 1 \text{ and } y \geq 0.$$

Hence, we shall use plus in front of the (real) square root, so we get after another rearrangement that

$$e^y = x + \sqrt{x^2 - 1} \geq 1, \qquad x \geq 1 \text{ and } y \geq 0,$$

thus

$$y = \operatorname{Arcosh} x = \ln\left(x + \sqrt{x^2 - 1}\right), \qquad x \geq 1.$$

Remark 2.3 We have above used that if we had chosen the other possibility,

$$e^y = x - \sqrt{x^2-1} < 1,$$

then $y < 0$, which is not possible. $\Diamond$

The derivative of Arcosh x. By differentiation we get for $x > 1$,

$$\begin{aligned}\frac{d}{dx}\,\mathrm{Arcosh}\,x &= \frac{d}{dx}\ln\left(x+\sqrt{x^2-1}\right) = \frac{1}{x+\sqrt{x^2-1}}\left(1+\frac{x}{\sqrt{x^2-1}}\right)\\ &= \frac{1}{x+\sqrt{x^2-1}}\cdot\frac{\sqrt{x^2-1}+x}{\sqrt{x^2-1}} = \frac{1}{\sqrt{x^2-1}}.\end{aligned}$$

Hence

$$(40)\quad \begin{cases} \mathrm{Arcosh}\,x = \ln\left(x+\sqrt{x^2-1}\right), & x \ge 1,\\[2mm] \dfrac{d}{dx}\,\mathrm{Arcosh}\,x = \dfrac{1}{\sqrt{x^2-1}}, & x > 1,\\[2mm] \cosh(\mathrm{Arcosh}\,x) = \cosh\left(\ln\left(x+\sqrt{x^2-1}\right)\right) = x, & x \ge 1.\end{cases}$$

The inverse of $y = \sinh x$. By using the same method as for $y = \cosh x$ we obtain

$$(41)\quad \begin{cases} \operatorname{Arsinh} x = \ln\left(x + \sqrt{x^2+1}\right), & x \in \mathbb{R}, \\[2mm] \dfrac{d}{dx}\operatorname{Arsinh} x = \dfrac{1}{\sqrt{x^2+1}}, & x \in \mathbb{R}, \\[2mm] \sinh(\operatorname{Arsinh} x) = \sinh\left(\ln\left(x+\sqrt{x^2+1}\right)\right) = x, & x \in \mathbb{R}. \end{cases}$$

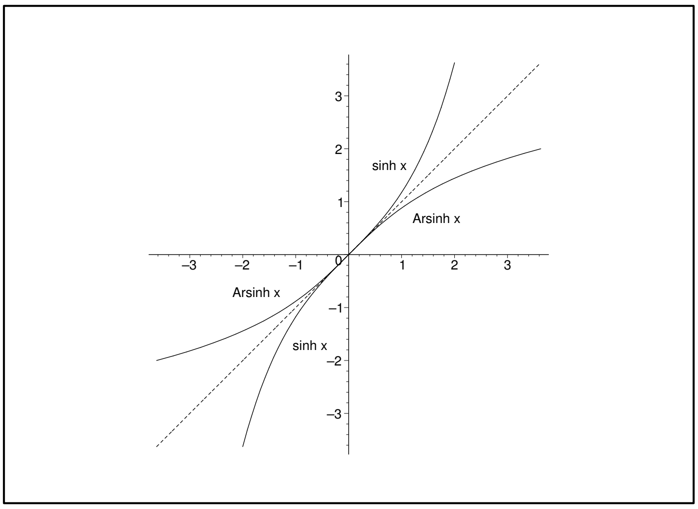

Figure 14: The graphs of $y = \sinh x$ and of $y = \operatorname{Arsinh} x = \ln\left(x + \sqrt{x^2+1}\right)$, $x \in \mathbb{R}$.

The inverse of $y = \tanh x$. When we solve the equation

$$x = \tanh y = \frac{e^{2y}-1}{e^{2y}+1}, \qquad x \in \,]-1,1[,$$

with respect to y we obtain

$$(42)\quad \begin{cases} \operatorname{Artanh} x = \dfrac{1}{2}\ln\left(\dfrac{1+x}{1-x}\right), & x \in \,]-1,1[, \\[2mm] \dfrac{d}{dx}\operatorname{Artanh} x = \dfrac{1}{1-x^2}, & x \in \,]-1,1[, \\[2mm] \tanh(\operatorname{Artanh} x) = \tanh\left(\dfrac{1}{2}\ln\left(\dfrac{1+x}{1-x}\right)\right) = x, & x \in \,]-1,1[. \end{cases}$$

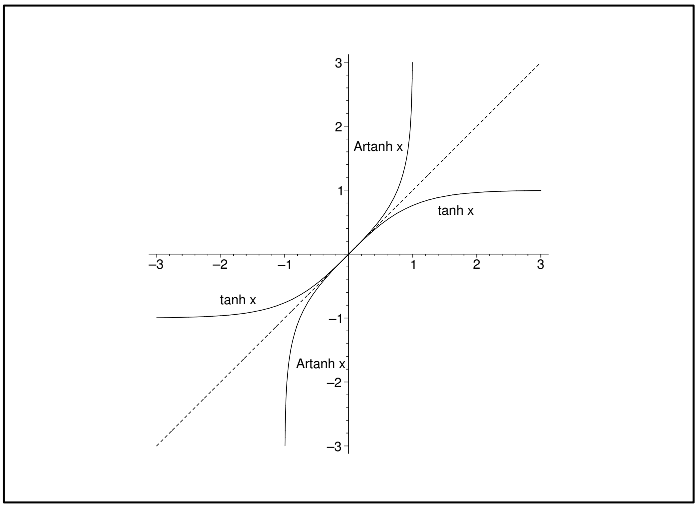

Figure 15: The graphs of $y = \tanh x$, $x \in \mathbb{R}$, and of

$$y = \operatorname{Artanh} x = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right), \qquad -1 < x < 1.$$

The inverse function of $y = \coth x$. When we solve the equation

$$x = \coth y = \frac{e^{2y}+1}{e^{2y}-1}, \qquad x \in \mathbb{R} \setminus [-1,1],$$

with respect to y we obtain

$$(43) \quad \begin{cases} \operatorname{Arcoth} x = \dfrac{1}{2}\ln\left(\dfrac{x+1}{x-1}\right), & x \notin [-1,1], \\[2ex] \dfrac{d}{dx}\operatorname{Arcoth} x = \dfrac{1}{1-x^2}, & x \notin [-1,1], \\[2ex] \coth(\operatorname{Arcoth} x) = \coth\left(\dfrac{1}{2}\ln\dfrac{x+1}{x-1}\right) = x, & x \notin [-1,1]. \end{cases}$$

Remark 2.4 The inverse functions of the trigonometric functions, cf. Section 1.9, can be represented by an arc. For that reason they are called *arcus functions*. Similarly, the inverse functions of the hyperbolic functions, considered here, can be represented by an area (they never are in the elementary textbooks). Thus, for historical reasons they are called *area functions*. $\Diamond$

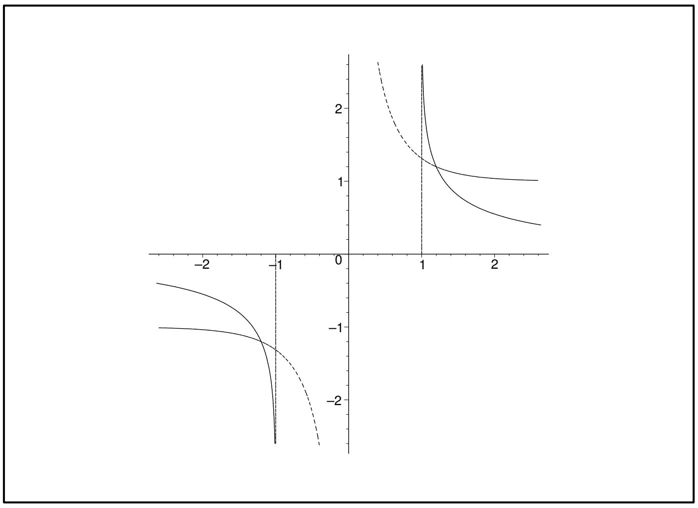

Figure 16: The graphs of $y = \coth x$, $x \neq 0$, (dotted) and of

$$y = \operatorname{Arcoth} x = \frac{1}{2} \ln\left(\frac{x+1}{x-1}\right), \qquad x \in \mathbb{R} \setminus [-1,1],$$

with its vertical asymptotes.

2.8 Arcus functions.

These functions are the inverse of the trigonometric functions, restricted to suitable intervals, due to the periodicity of the trigonometric functions.

The inverse functions of the trigonometric functions cannot be given an explicit expression as a real function composed of the logarithm and the square root, as in the case of the area functions. Therefore, for the time being the only method of describing them is to restrict the domain of the corresponding trigonometric functions and reflect the graphs of these restrictions in order to obtain the graphs of the inverse functions. The derivatives are found by means of the methods from Section 1.2.

The four arcus functions considered here are denoted by

$$\operatorname{Arccos} x, \quad \operatorname{Arcsin} x, \quad \operatorname{Arctan} x, \quad \operatorname{Arccot} x.$$

Other notations are

$$\begin{array}{llll} \arccos x, & \arcsin x, & \arctan x, & \operatorname{arccot} x, \\ \cos^{-1} x, & \sin^{-1} x, & \tan^{-1} x, & \cot^{-1}. \end{array}$$

It is recommended to avoid the expressions in the last line, since e.g. $\cos^{-1} x$ may be confused with $1/\cos x$, etc.

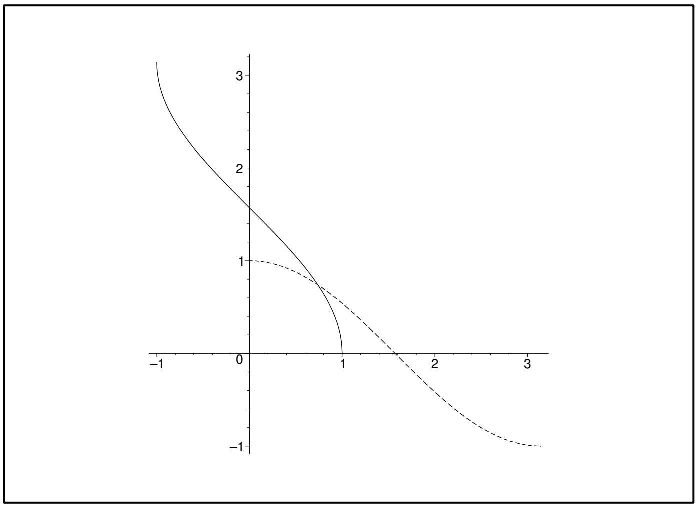

Figure 17: The graphs of $y = \text{Arccos}\, x$, $x \in [-1, 1]$, and $y = \cos x$, $x \in [0, \pi]$, (dotted).

The function Arccos x. When the graph of $y = \cos x$ is considered, we see that it is decreasing in the (maximum) interval $[0, \pi]$, so the inverse function exists for this particular restriction.

It follows that the function

$$y = \text{Arccos}\, x \in [0, \pi], \qquad x \in [-1, 1],$$

is only defined by reflecting the graph of $y = \cos x$, $x \in [0, \pi]$, and by the relation

$$\cos(\text{Arccos}\, x) = x, \qquad x \in [-1, 1].$$

Using the method of differentiation of an inverse function, given in Section 2.2, we obtain

$$\frac{d}{dx}\,\text{Arccos}\, x = \frac{1}{-\sin(\text{Arccos}\, x)} = -\frac{1}{\sqrt{1-\cos^2(\text{Arccos}\, x)}} = -\frac{1}{\sqrt{1-x^2}}.$$

Summary.

$$(44) \quad \begin{cases} y = \text{Arccos}\, x \in [0, \pi], & \text{for } x \in [-1, 1], \\[2mm] \cos(\text{Arccos}\, x) = x, & \text{for } x \in [-1, 1], \\[2mm] \dfrac{d}{dx}\,\text{Arccos}\, x = -\dfrac{1}{\sqrt{1-x^2}}, & \text{for } x \in]-1, 1[. \end{cases}$$

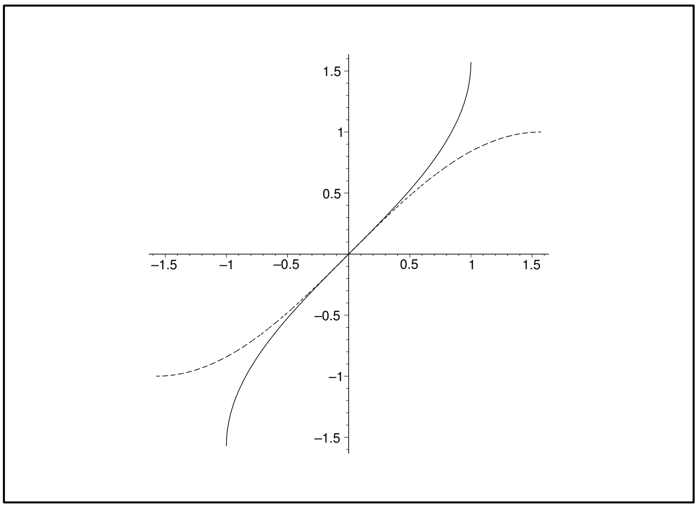

Figure 18: The graphs of $y = \text{Arcsin}\, x$, $x \in [-1, 1]$, and $y = \sin x$, $x \in [-\pi/2, \pi/2]$, (dotted).

The function Arcsin x. Among many possible restrictions of $y = \sin x$ we choose

$$y = \sin x, \qquad x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

In this case we have

$$(45)\quad \begin{cases} y = \operatorname{Arcsin} x \in \left[-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right], & \text{for } x \in [-1,1], \\[2mm] \sin(\operatorname{Arcsin} x) = x, & \text{for } x \in [-1,1], \\[2mm] \dfrac{d}{dx} \operatorname{Arcsin} x = \dfrac{1}{\sqrt{1-x^2}}, & \text{for } x \in]-1,1[. \end{cases}$$

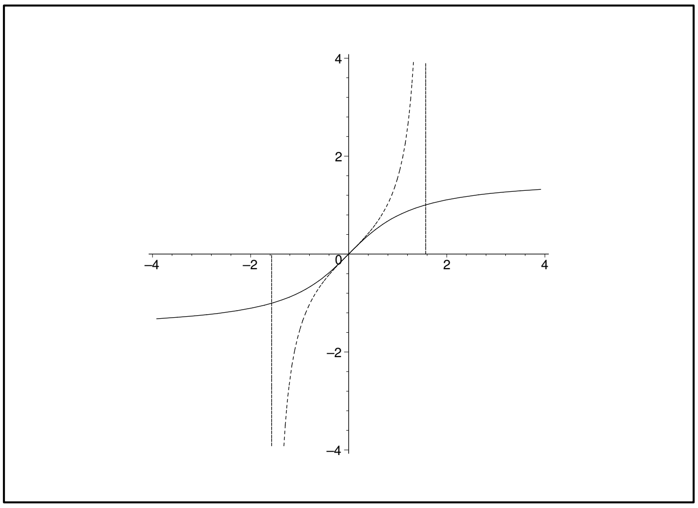

Figure 19: The graphs of $y = \operatorname{Arctan} x$, $x \in \mathbb{R}$, and $y = \tan x$, $x \in]-\pi/2, \pi/2[$ (dotted with dotted asymptotes).

The function Arctan x. Among many possibilities we choose the restriction

$$y = \tan x \in \mathbb{R} \qquad \text{for } x \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[.$$

In this case we have

$$(46)\quad \begin{cases} y = \arctan x \in \left]-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right[, & \text{for } x \in \mathbb{R}, \\[2mm] \tan(\operatorname{Arctan} x) = x, & \text{for } x \in \mathbb{R}, \\[2mm] \dfrac{d}{dx} \operatorname{Arctan} x = \dfrac{1}{1+x^2}, & \text{for } x \in \mathbb{R}. \end{cases}$$

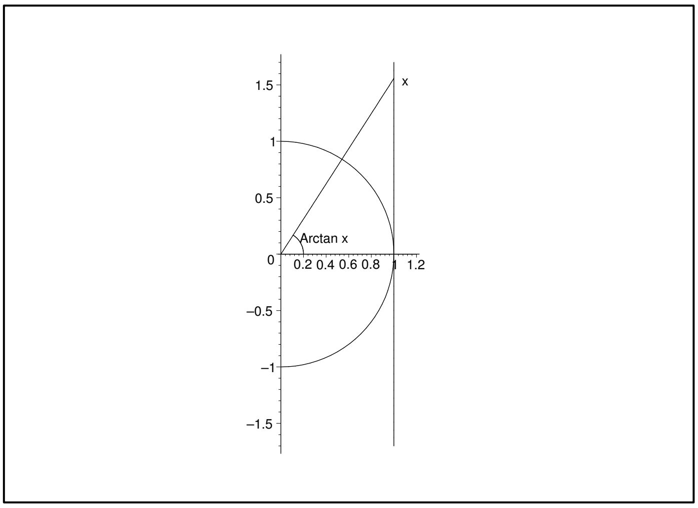

Figure 20: If x is the coordinate on the axis through $(1,0$, parallel with the Y-axis, then Arctan x is the angle between the positive X-axis and the vector $(1,x)$.

The function Arccot x. Among many possible restrictions we choose

$$y = \cot x \in \mathbb{R}, \qquad \text{for } x \in]0,\pi[.$$

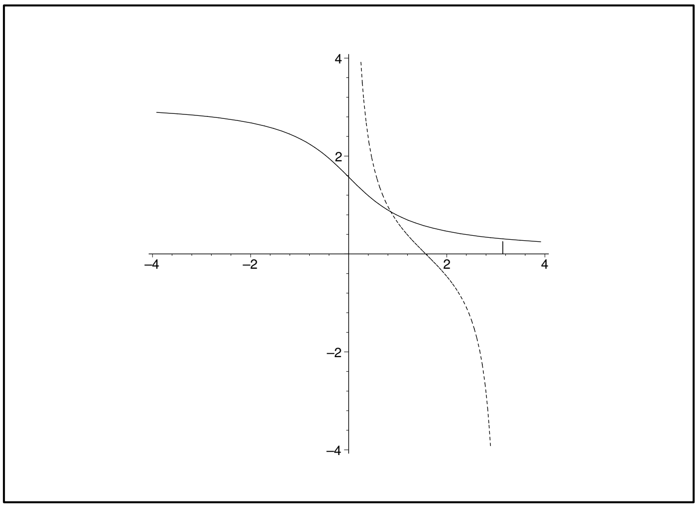

Figure 21: The graphs of $y = \operatorname{Arccot} x$, $x \in \mathbb{R}$, and $y = \cot x$, $x \in]0,\pi[$ (dotted).

In this case we have

$$(47)\quad \begin{cases} y = \operatorname{Arccot} x \in \,]0,\pi[, & \text{for } x \in \mathbb{R}, \\ \cot(\operatorname{Arccot} x) = x, & \text{for } x \in \mathbb{R}, \\ \dfrac{d}{dx}\operatorname{Arccot} x = -\dfrac{1}{1+x^2}, & \text{for } x \in \mathbb{R}. \end{cases}$$

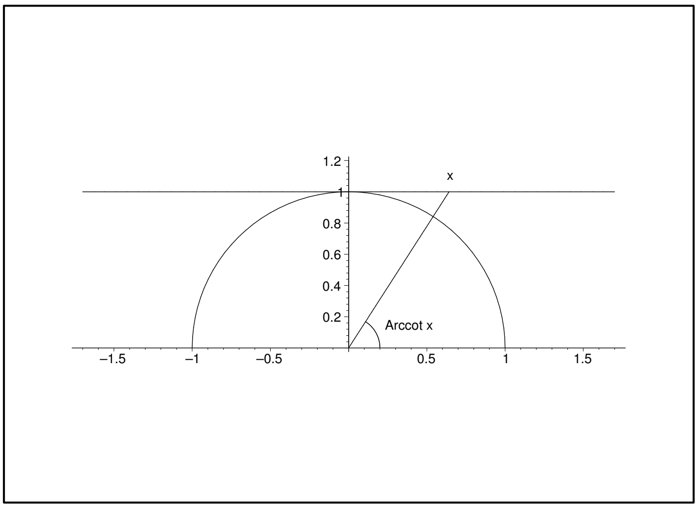

Figure 22: If x is the coordinate on the axis through $(0, 1)$ parallel with the X-axis, then Arccot x is the angle between the positive Xaxis and the vector $(x, 1)$.

Relations. Just like the trigonometric functions the arcus functions have been investigated for centuries. Many relations have been found, of which we only show the simplest ones; cf. also the appendix.

By combining the derivatives of (44) and (45), or (46) and (47), we get

$$\frac{d}{dx}\{\operatorname{Arccos} x + \operatorname{Arcsin} x\} = -\frac{1}{\sqrt{1-x^2}} + \frac{1}{\sqrt{1-x^2}} = 0 \qquad \text{for } x \in \,]-1,1[,$$

and

$$\frac{d}{dx}\{\operatorname{Arctan} x + \operatorname{Arccot} x\} = \frac{1}{1+x^2} - \frac{1}{1+x^2} = 0, \qquad \text{for } x \in \mathbb{R}.$$

Since the functions are of class C^1 with the derivative 0, they must be constants. These constants are found by choosing e.g. $x = 0$. Hence, by a trivial continuous extension in the first case,

$$\operatorname{Arccos} x + \operatorname{Arcsin} x = \frac{\pi}{2} \qquad \text{for } x \in [-1,1],$$

and

$$\operatorname{Arctan} x + \operatorname{Arccot} x = \frac{\pi}{2} \qquad \text{for } x \in \mathbb{R}.$$

2.9 Magnitude of functions.

In many cases we have to compare two functions, which both tend to infinity when $x \to +\infty$ or $x \to -\infty$. To ease matters in the following chapters some results are given here after our introduction of our most common functions in this chapter. Notice that even if we only consider four classes of functions, *logarithms, power functions, exponentials* and *faculty functions*, we get a lot of information from the fact that these four types of functions all tend to $+\infty$ for $x \to +\infty$, but with a different speed. Thus we can establish a crude hierarchy of the functions according to their magnitudes in infinity.

We shall prove the following very important theorem.

Theorem 2.1 *The logarithms, the power functions, the exponentials and the faculty functions form a hierarchy concerning their magnitudes when $x \to +\infty$ in the following sense.*

1) A power function will always dominate a logarithm, when $x \to +\infty$. More precisely, for any constants α, $\beta > 0$,

$$\frac{(\ln x)^\alpha}{x^\beta} \to 0 \qquad \text{for } x \to +\infty.$$

For completion we add that also

$$x^\beta |\ln x|^\alpha \to 0 \qquad \text{for } x \to 0+\,.$$

2) An exponential will always dominate a power function, when $x \to +\infty$. More precisely, for any constants $\alpha > 0$ and $a > 1$,

$$\frac{x^\alpha}{a^x} \to 0 \qquad \textit{for } x \to +\infty.$$

3) A faculty function will always dominate an exponential for $n \to +\infty$, $n \in \mathbb{N}$. More precisely, if $a > 1$ is a constant, then

$$\frac{a^n}{n!} \to 0 \qquad \textit{for } n \in \mathbb{N} \textit{ and } n \to +\infty.$$

Remark 2.5 Item 3) is only rarely mentioned in a first course in Calculus, in spite of the fact that it is also very important. Furthermore, for the time being the faculty function $n!$ is only defined for $n \in \mathbb{N}_0$. In later courses it will be extended to the gamma function $\Gamma(z)$, which is defined and continuous even for $z \in \mathbb{C} \setminus \{\{0\} \cup \mathbb{Z}_-\}$. The gamma function is defined such that $\Gamma(n+1) = n!$ for $n \in \mathbb{N}_0$. Once this function has been introduced, 3) can be extended to

$$\frac{a^x}{\Gamma(x+1)} \to 0 \quad \text{or} \quad \frac{a^x}{\Gamma(x)} \to 0 \qquad \text{for } x \to +\infty. \qquad \Diamond$$

Proof of Theorem 2.1. We shall first prove 2).

First note that since $e^x > 1$ for $x > 0$ we have

$$e^x = 1 + \int_0^x e^t \, dt > 1 + \int_0^x 1 \, dt = 1 + x.$$

When this estimate is iterated we get in the next step that

$$e^x = 1 + \int_0^x e^t \, dt > 1 + \int_0^x (1+t) \, dt = 1 + \frac{x}{1!} + \frac{x^2}{2!}.$$

Then it follows easily by induction that

$$(48) \quad e^x = 1 + \int_0^x e^t \, dt > 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}, \qquad x > 0, \qquad \text{for every } n \in \mathbb{N},$$

Let $\alpha > 0$ be given and choose $n \in \mathbb{N}$, such that $\alpha < n$. Since it follows trivially from (48) that in general

$$e^x > \frac{x^n}{n!} \qquad \text{for all } x > 0 \text{ and all } n \in \mathbb{N},$$

we see that we get

$$(49) \quad \frac{x^\alpha}{e^x} < n! \, \frac{x^\alpha}{x^n} = \frac{n!}{x^{n-\alpha}} \to 0 \qquad \text{for } n > \alpha \text{ fixed and } x \to +\infty.$$

This proves 2) in the special case of $a = e$.

If $a > 1$, we have $a^x = \exp(x \ln a)$, where $\ln a > 0$. If we write $y = x \ln a$, we get according to (49),

$$\frac{x^\alpha}{a^x} = \frac{1}{(\ln a)^\alpha} \cdot \frac{y^\alpha}{e^y} \to 0 \qquad \text{for } y \to +\infty, \text{ i.e. for } x \to +\infty,$$

and 2) is proved.

Next we prove 1).
Let $t = e^x$, i.e. $x = \ln t$, and let α, $\beta > 0$ be positive constants. Then

$$(50) \quad \frac{(\ln t)^\alpha}{t^\beta} = \frac{x^\alpha}{(e^x)^\beta} = \frac{x^\alpha}{(e^\beta)^x} = \frac{x^\alpha}{a^x} \to 0 \qquad x \to +\infty,$$

according to 2), because $a := e^\beta > 1$, when $\beta > 0$.

Furthermore, if we put $t = \dfrac{1}{s}$ in (50), then $s \to 0+$ when $t \to +\infty$, hence

$$s^\beta |\ln s|^\alpha = \frac{|\ln t|^\alpha}{t^\beta} = \frac{(\ln t)^\alpha}{t^\beta} \to 0 \qquad \text{for } s \to 0+,$$

and 1) is proved.

Finally, 3) is easy to prove.
Let $a > 1$ and choose $n \in \mathbb{N}$, such that e.g. $n > 2a$. Then for every $p \in \mathbb{N}$,

$$0 < \frac{a^{n+p}}{(n+p)!} = \frac{a^n}{n!} \cdot \frac{a}{n+1} \cdots \frac{a}{n+p} < \frac{1}{2^p} \cdot \frac{a^n}{n!} \to 0 \qquad \text{for } p \to +\infty,$$

since $n \in \mathbb{N}$ is fixed. Thus 3) is proved. $\square$

Remark 2.6 It is natural to include $\sinh x$ and $\cosh x$ in the exponential class, and polynomials and roots of polynomials in the power function class, etc.. $\Diamond$

Remark 2.7 In the notation from Chapter 6 we can write

$$\frac{(\ln x)^\alpha}{x^\beta} = \varepsilon\left(\frac{1}{x}\right) \qquad \text{for } \alpha \in \mathbb{R} \text{ and } \beta \in \mathbb{R}_+ \text{ and } x \to +\infty,$$

$$\frac{x^\alpha}{a^x} = \varepsilon\left(\frac{1}{x}\right) \qquad \text{for } \alpha \in \mathbb{R} \text{ and } a > 1 \text{ and } x \to +\infty,$$

$$\frac{a^n}{n!} = \varepsilon\left(\frac{1}{n}\right) \qquad \text{for } a \in \mathbb{R} \text{ and } n \to +\infty, \quad n \in \mathbb{N}.$$

Notice that we in the first two statements can allow $\alpha \in \mathbb{R}$, and that we in the last statement can allow any $a \in \mathbb{R}$. $\Diamond$

3 Differentiation

3.1 Introduction.

At this stage of learning mathematics one considers differentiation as a skilled trade in opposition to integration, which is an art. The reason is that we still only consider a rather narrow class of functions, in which differentiation is a simple operation, while one concerning integration may get some problems if one wants explicit results.

It should here be mentioned that when one leaves ordinary calculus and starts on more advanced mathematical analysis, then it turns up that the differentiation operator is a real villain, while the integration operator is smoothing out all the bad things. Hence, in Mathematical Analysis one should always try to write formulæ in terms of integration instead of in terms of differentiation, whenever possible, while one in Calculus would prefer differentiation instead of integration.

Since this is Calculus, we shall stick to the traditional view of differentiation as the easy operator, and yet we cannot totally avoid the hidden complications mentioned above.

3.2 Definition and geometrical interpretation.

A function $f : I \in \mathbb{R}$ is said to be differentiable at an inner point $x_0 \in I$, if the limit of the difference quotient

(51) $$f'(x_0) := \lim_{\substack{x \to x_0 \\ x \in I \setminus \{x_0\}}} \frac{f(x) - f(x_0)}{x - x_0}$$

exists. The limit is denoted by

$$f'(x_0), \qquad \frac{df}{dx}(x_0), \qquad Df(x_0),$$

or similarly. It is called the *derivative of the function f at the inner point $x_0 \in I$.*

The derivative $f'(x_0)$ of the function f at the point x_0 is interpreted as the slope of the tangent of the curve $y = f(x)$ at the point $(x_0, f(x_0))$. Hence, the equation of the tangent is

(52) $$y = f(x_0) + f'(x_0) \cdot (x - x_0).$$

Since almost every approach of an application starts by a linearization of the underlying mathematical model, i.e. the curve $y = f(x)$ is replaced by the tangent (52) in the neighbourhood of $x = x_0$, it is seen that (52) is very important, and yet it is trivial at the same time.

It is obvious that the definition (51) can be used for any $x_0 \in I$, whenever the limit exists. If $x_0 \in I$ is a boundary point which is not an isolated point, we may talk about “half tangents”, or tangents from the left (right), whenever x_0 is an end point of an interval I.

Suppose that $f : T \to \mathbb{R}$ is differentiable at every point of a nonempty subset $I_0 \subseteqq I$. Then the derivative can be considered as a function

$$f_0' : I_0 \to \mathbb{R}.$$

If this derivative $f' : I_0 \to \mathbb{R}$ is continuous, we say that f is continuously differentiable on I_0, and we write $f \in C^1(I_0)$.

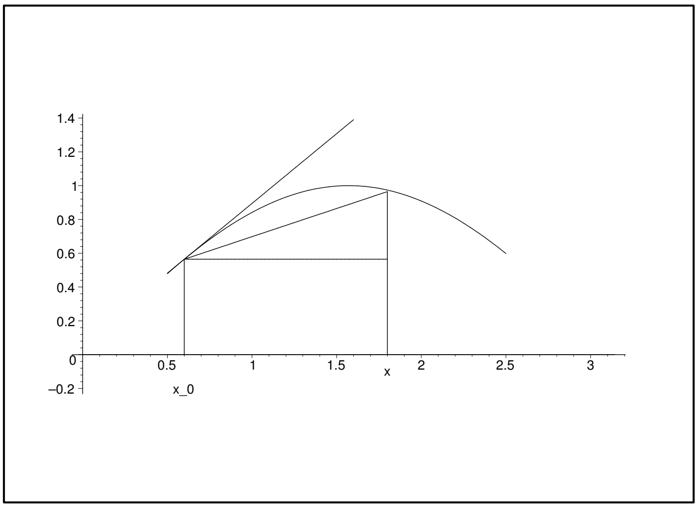

Figure 23: Geometrical interpretation of the derivative.

By induction, the class $C^n(I_0)$ is defined as the set of functions $f : I_0 \to \mathbb{R}$, for which the successive derivatives f, f', ..., $f^{(n)}$ up to order $n \in \mathbb{N}$ all exist and are continuous on the open nonempty set I_0.

If we define $C^0(I_0)$ as the set of all continuous functions on I_0, we obviously have the descending chain

$$C^0(I_0) \supset C^1(I_0) \supset C^2(I_0) \supset \cdots \supset C^n(I_0) \supset \cdots .$$

By taking the limit we finally get

$$C^\infty(I_0) = \bigcap_{n=0}^{\infty} C^n(I_0) = \lim_{n\to\infty} C^n(I_0),$$

so $C^\infty(I_0)$ is the set of functions, which are differentiable of any order, $f^{(n)}$, $n \in \mathbb{N}_0$, where each $f^{(n)}$ is continuous on I_0. Since e.g. the constants and the polynomials obviously all belong to $C^\infty(\mathbb{R})$, we see that $C^\infty(I_0)$ is not empty.

All the strange looking manoeuvres above are necessary, because of what we mentioned vaguely in the introduction, namely that differentiation introduced by the definition (51) has some very bad properties. The pedagogical problem at this stage of the development is that since one in the elementary Calculus usually starts with working in spaces of the type $C^\infty(I_0)$ (in fact in the smaller class of analytical functions which behave very nicely) where the differentiation given by (51) is straightforward, the student does not immediately see why the tedious accuracy above really is necessary.

We must therefore for historical reasons mention here that it came as a shock for the mathematical world, when Karl Weierstraß at the end of the nineteenth century constructed a function $f \in C^0(\mathbb{R})$ which is not differentiable at any point in $\mathbb{R}$!

One has ever since (like here) always been keen on defining the spaces C^0, C^1, C^n and C^∞ properly. And then in most cases it is forgotten to explain why it was necessary.

Any author of books in Calculus has been in this dilemma, because Weierstraß's function cannot be explained at this early stage of the development of Calculus. One needs at least knowledge of Fourier series and uniform convergence before it is possible, and even when these are known, the construction is not an easy one.

3.3 A catalogue of known derivatives.

Some explicit derivatives are already known from high school and Chapter 2. They are listed once more here, because they form the building stones of more advanced differentiations, and then it would be convenient to keep them all at the same place. They ought to be mastered by the students, so the differentiation formulæ in the applications should not be checked all the time.

Power functions. If $f(x) = x^\alpha$, $x \in \mathbb{R}_+$ and $\alpha \in \mathbb{R}$, then

$$(53)\quad f'(x) = \frac{d}{dx}\left(x^\alpha\right) = \alpha \cdot x^{\alpha-1}, \qquad \text{for } x \in \mathbb{R}_+.$$

Notice that this result is unchanged, if f can be extended to a larger open set. We must only be careful with the case $x = 0$, because we are never allowed to divide by zero. E.g. $\sqrt[3]{x}$ is defined for all $x \in \mathbb{R}$, but its derivative $\dfrac{1}{3}\, x^{-2/3}$ is not.

Logarithms. It follows from Chapter 2 that

(54) $$\frac{d}{dx}\ln x = \frac{1}{x} \qquad \text{for } x \in \mathbb{R}_+,$$

and

(55) $$\frac{d}{dx}\log_g(x) = \frac{d}{dx}\left\{\frac{\ln x}{\ln g}\right\} = \frac{1}{x \cdot \ln g}, \qquad x \in \mathbb{R}_+, \text{ where } g \in \mathbb{R}_+ \setminus \{1\}.$$

Exponentials. We have seen that

(56) $$\frac{d}{dx}\exp(x) = \frac{d}{dx}e^x = e^x = \exp(x), \qquad \text{for } x \in \mathbb{R},$$

so $\exp(x)$ is not changed by a differentiation with respect to x.

We have more generally that

(57) $$\frac{d}{dx}(a^x) = \frac{d}{dx}\exp(x \cdot \ln a) = \ln a \cdot a^x, \qquad \text{for } x \in \mathbb{R} \text{ and } a \in \mathbb{R}_+.$$

Trigonometric functions. It is known from high school that

(58) $$\frac{d}{dx}\sin x = \cos x, \qquad \frac{d}{dx}\cos x = -\sin x, \qquad \text{for } x \in \mathbb{R},$$

and

(59) $$\begin{cases} \dfrac{d}{dx}\tan z = \dfrac{1}{\cos^2 x} = 1 + \tan^2 x, & x \neq \dfrac{\pi}{2} + p\pi, \qquad p \in \mathbb{Z}, \\[2ex] \dfrac{d}{dx}\cot z = -\dfrac{1}{\sin^2 x} = -\left(1 + \cot^2 x\right), & x \neq p\pi, \qquad p \in \mathbb{Z}. \end{cases}$$

Hyperbolic functions.

(60) $$\frac{d}{dx}\sinh x = \cosh x, \qquad \frac{d}{dx}\cosh x = \sinh x, \qquad x \in \mathbb{R},$$

and

(61) $$\begin{cases} \dfrac{d}{dx}\tanh x = \dfrac{1}{\cosh^2 x} = 1 - \tanh^2 x, & x \in \mathbb{R}, \\[2ex] \dfrac{d}{dx}\coth x = -\dfrac{1}{\sinh^2 x} = 1 - \coth^2 x, & x \in \mathbb{R} \setminus \{0\}. \end{cases}$$

Inverse trigonometric functions. The derivatives were found in Chapter 2, so we just quote

(62) $$\frac{d}{dx}\text{Arcsin}\, x = \frac{1}{\sqrt{1-x^2}}, \qquad \frac{d}{dx}\text{Arccos}\, x = -\frac{1}{\sqrt{1-x^2}}, \qquad x \in\,]-1,1[,$$

and

(63) $$\frac{d}{dx}\text{Arctan}\, x = \frac{1}{1+x^2}, \qquad \frac{d}{dx}\text{Arccot}\, x = -\frac{1}{1+x^2}, \qquad x \in \mathbb{R}.$$

Inverse hyperbolic functions. Their derivatives were found in Chapter 2:

$$(64)\quad \begin{cases} \dfrac{d}{dx}\,\mathrm{Arsinh}\,x = \dfrac{1}{\sqrt{x^2+1}}, & x \in \mathbb{R}, \\[2mm] \dfrac{d}{dx}\,\mathrm{Arcosh}\,x = \dfrac{1}{\sqrt{x^2-1}}, & x \in \,]1,+\infty[, \\[2mm] \dfrac{d}{dx}\,\mathrm{Artanh}\,x = \dfrac{1}{1-x^2}, & x \in \,]-1,1[, \\[2mm] \dfrac{d}{dx}\,\mathrm{Arcoth}\,x = \dfrac{1}{1-x^2}, & x \in \mathbb{R} \setminus [-1,1]. \end{cases}$$

Notice in particular that the derivatives of Artanh x and Arcoth x formally are identical in structure, but since they are not defined in the same set, they are nevertheless different.

3.4 The simple rules of calculation.

We assume everywhere in this section that f and g are continuously differentiable, so this assumption will not always be repeated.

Differentiation is a linear operator. If $a, b \in \mathbb{R}$ are constants, and f, $g \in C^1$, then $a \cdot f + b \cdot g \in C^1$ and

$$(65)\quad \frac{d}{dx}\{a \cdot f(x) + b \cdot g(x)\} = a \cdot f'(x) + b \cdot g'(x).$$

Differentiation of a product. If f, $g \in C^1$, then $f \cdot g \in C^1$, and

$$(66)\quad \frac{d}{dx}\{f(x) \cdot g(x)\} = f'(x) \cdot g(x) + f(x) \cdot g'(x).$$

Differentiation of a reciprocal function. If $g \in C^1$ and $g(x) \neq 0$ in the domain, then $\dfrac{1}{g} \in C^1$, and

$$(67)\quad \frac{d}{dx}\left\{\frac{1}{g(x)}\right\} = -\frac{g'(x)}{g(x)^2}.$$

Differentiation of a quotient. Assume that f, $g \in C^1(I)$, where $g(x) \neq 0$ for all $x \in I$. Then also $\dfrac{f}{g} \in C^1(I)$, and its derivative can be calculated in two ways (of course both leading to the same result):

1) From high school Calculus we have the formula

$$(68)\quad \frac{d}{dx}\left\{\frac{f(x)}{g(x)}\right\} = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}, \qquad x \in I.$$

2) Formula (68) often gives large and complicated calculations. It is worth to pay attention to the following *alternative* method, where we use the rule of differentiation of a product followed by the differentiation of a reciprocal function,

$$(69)\quad \frac{d}{dx}\left\{\frac{f(x)}{g(x)}\right\} = \frac{d}{dx}\left\{f(x) \cdot \frac{1}{g(x)}\right\} = \frac{f'(x)}{g(x)} - f(x) \cdot \frac{g'(x)}{g(x)^2}.$$

It is easily seen that (68) and (69) are identical and that (69) in general is simpler than (68), because we have cancelled the common factor $g(x)$ in the first of the two fractions.

3.5 Differentiation of composite functions.

Let $f : I \to J$ and $g : J \to \mathbb{R}$ be C^1-functions. The the composite function

$$g \circ f : I \to \mathbb{R}, \qquad (g \circ f)(x) := g(f(x)),$$

is also a C^1-function, and its derivative is given by

$$(70) \quad \frac{d}{dx}\,(g \circ f)(x) = g'(f(x)) \cdot f'(x).$$

This result has already been used several times, and it will be used over and over again, often without being conscious of that it is this rule we are using.

An *alternative* description of (70) is the following incorrect use of the symbols, which nevertheless leads to the right result:

If we put $y = f(x)$ and $z = g(y)$, then

$$(71) \quad \frac{d(g \circ f)}{dx} = \frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx} = g'(y) \cdot f'(x) = g'(f(x)) \cdot f'(x).$$

This can be expressed by saying that z is differentiated with respect to x through y. This is also called the *chain rule*.

3.6 Differentiation of an implicit given function.

In this section we are forced to anticipate some results on functions in two variables. This is not the usual way of exposing Calculus, but I hope that this procedure may dissolve some of the difficulties in the understanding of what is going on without hopefully introducing too many new problems of understanding.

Let $\Omega \subseteq \mathbb{R}^2$ be a nonempty open set, and let $F : \Omega \to \mathbb{R}$ be a continuous function. Assume that $F(\cdot, y)$ is of class C^1 in x for every fixed y for which $(x, y) \in \Omega$, and that $F(x, \cdot)$ is of class C^1 in y for every fixed x for which $(x, y) \in \Omega$.

We say that $F \in C^1(\Omega)$, and the derivative of $F(\cdot, y)$ with respect to x for any fixed y is denoted by

$$\frac{\partial F}{\partial x}, \qquad \frac{\partial F}{\partial x}(x, y), \qquad \text{or} \qquad F'_x(x, y).$$

Similarly we introduce

$$\frac{\partial F}{\partial y}, \qquad \frac{\partial F}{\partial y}(x, y), \qquad \text{or} \qquad F'_y(x, y),$$

and $F'_x(x, y)$ and $F'_y(x, y)$ are called the *partial derivatives* of $F(x, y)$ with respect to x and y.

Let $(x_0, y_0) \in \Omega$ be a point such that

$$(72) \quad F(x_0, y_0) \qquad \text{and} \qquad F'_y(x_0, y_0) \neq 0, \qquad F \in C^1(\Omega).$$

Then it can be proved that the equation

$$F(x,y)=0$$

in a neighbourhood of $x = x_0$, e.g. given by $|x - x_0| < \varepsilon$, defines a unique C^1-function $y = y(x)$, such that

(73) $F(x,y(x)) = 0, \qquad \text{for } |x-x_0|<\varepsilon.$

This result is the content of the simplest form of the *theorem of implicit given functions*.

The derivative $y'(x)$ of this implicitly given function can be bound by "implicit differentiation" of (73), i.e. by use of the *chain rule*,

(74) $\dfrac{\partial F}{\partial x}\cdot\dfrac{dx}{dx}+\dfrac{\partial F}{\partial y}\cdot\dfrac{dy}{dx}=0,$

which we here take for granted. (It is dealt with this in Calculus 2).

In the present specific case (74) is written

(75) $F'_x(x,y(x)) + F'_y(x,y(x))\cdot y'(x) = 0.$

Since $F_y'(x_0, y_0) = F_y'(x_0, y(x_0)) \neq 0$, we also have by continuity that for some sufficiently small $\varepsilon > 0$,

$$F_y'(x, y(x)) \neq 0 \qquad \text{for } |x - x_0| < \varepsilon.$$

Hence we get from (75) that

$$(76)\quad y'(x) = -\frac{F_x'(x, y(x))}{F_y'(x, y(x))}, \qquad |x - x_0| < \varepsilon.$$

It follows immediately from (76) why we have to assume that $F_y'(x, y(x)) \neq 0$.

this method can be extended. If $F \in C^2(\Omega)$, while the other assumptions are unchanged, then $y = y(x)$ is also defined and of class C^2 for $|x - x_0| < \varepsilon$, where $\varepsilon > 0$ is chosen appropriately. Then $y''(x)$ is found by differentiating either (75) or (76) with respect to x.

It is obvious how we proceed when $F \in C^n(\Omega)$.

3.7 Differentiation of an inverse function.

This is a special case of Section 3.6.

Let $f : I \to J$ be one-to-one and strictly monotone and of class C^1. There exists a unique *inverse function* $f : J \to I$, which is also of class C^1, such that

$$(f \circ g)(x) = f(g(x)) = x \qquad \text{for } x \in J.$$

In this case we choose

$$(77)\quad F(x, y) = f(y) - x = 0 \qquad \text{and} \qquad y = g(x),$$

so (75) can now be written as

$$f'(g(x)) \cdot g'(x) = 1,$$

from which

$$(78)\quad g'(x) = \frac{1}{f'(g(x))},$$

which was also found in Chapter 2.

If furthermore f is of class C^2, then of course

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

We proceed in this way to find the higher order derivatives if f is of class C^n.

4 Integration

4.1 Introduction.

At this stage of the student's education, integration is considered as an art, while differentiation only is a skilled trade. The reason is that the goal here is to find the exact expression of the integration, resp. differentiation of a given nice function.

It should be mentioned once more that in more advanced applications one often has to abandon the idea of finding the exact result of an integration, e.g. when the function under consideration is represented by a Fourier series (cf. Calculus 3). When we do this we have to concentrate on the properties of the two operators, and then it turns up that integration is much nicer (it is "smoothing") than differentiation where one typically gets some ugly cusps.

However, for the time being we must acknowledge that it usually is far more difficult to find the exact result of integration of a function than of differentiation.

Apart from an additional arbitrary constant, which always occurs if one does not specify the limits of integration, we shall here consider integration as the inverse operator of differentiation. One also calls it the *antiderivative* as opposed to the *derivative*. This means that if we find

(79) $F(x) = \int f(x)\, dx,$

then we can always *check our calculations by differentiating the result*,

(80) $F'(x) = f(x).$

This is trivial, and yet this principle can from time to time be used to ease the calculations as an alternative to some horrible standard solution formula. Sometimes one can in fact guess from the look of $f(x)$ how the structure of the antiderivative should be, e.g.

(81) $F(x) = \int f(x)\, dx = a_1\varphi_1(x) + \cdots + a_n\varphi_n(x),$

where the functions φ_i are known, while the constants $a_1, \dots, a_n$ still need to be specified. According to (80) we must necessarily have

(82) $F'(x) = f(x) = a_1\varphi_1'(x) + \cdots + a_n\varphi_n'(x),$

which typically produces n linear equations in the n unknown constants $a_1, \dots, a_n$. If the functions $\varphi_1, \dots, \varphi_n$ are linearly independent, it can be shown that the solution is unique.

Some of the earlier problems with finding an exact integral can now be circumvented by using advanced "pocket calculators" like TI-92 or newer models, or computer programs like MAPLE and MATHEMATICA. These must not, however, be a pretext for doing nothing, because none of them are without errors or shortcomings. By chance I once came across an error in one of the integration formulæ of TI-92, which I reported back to the company. This particular error should now have been removed in the newer models, but this is no guarantee for that no other error exists.

Concerning MAPLE and MATHEMATICA, these two programs have the tendency occasionally to give the result as a sum of hypergeometric functions, and since these are not known in the elementary Calculus, such a result is not of much help.

Furthermore, I have also been able to construct a family of functions $f_\alpha(x)$ where I could find the exact value of their definite integrals from 0 to $\pi/2$, while MAPLE and MATHEMATICA only gave decimal fractions. For some particular values of α these programs would even give complex results in spite of the fact that the integrand was nonnegative everywhere.

Example 4.1 *Find the exact value of*

$$(83)\quad \int_0^{\pi/2} f_\alpha(x)\,dx = \int_0^{\pi/2} \frac{dx}{1+\tan^\alpha x} \qquad \textit{for every } \alpha \in \mathbb{R}.$$

By using high school Calculus it is easy to find the result $\pi/4$, when $\alpha = 0$, 1, 2. When $\alpha = \frac{1}{2}$, or any number of the type $n + \frac{1}{2}$, where $n \in \mathbb{N}_0$, an earlier version of MAPLE gave a complex number, which obviously is wrong. And when $\alpha = \sqrt{2}$, both MAPLE and MATHEMATICA gave a decimal fraction as their result. Nevertheless it is easy to find the exact value of (83) for every $\alpha \in \mathbb{R}$.

First notice that the integrand of (83) lies in the interval $[0,1]$ and that it has a continuous extension to $[0,\pi/2]$, no matter the choice of $\alpha \in \mathbb{R}$. Therefore, the integral (83) exists for all $\alpha \in \mathbb{R}$, and its value lies in the interval $[0,\pi/2]$.

By the change of variable, $x \mapsto \frac{\pi}{2} - x$, we see that $\tan x$ is transformed into $\cot x$. Hence,

$$\int_0^{\pi/2} \frac{1}{1+\tan^\alpha x}\,dx = \int_0^{\pi/2} \frac{1}{1+\cot^\alpha x}\,dx = \int_0^{\pi/2} \frac{\tan^\alpha x}{1+\tan^\alpha x}\,dx.$$

The common value is therefore given by

$$\int_0^{\pi/2} \frac{1}{1+\tan^\alpha x}\,dx = \frac{1}{2}\left\{\int_0^{\pi/2} \frac{1}{1+\tan^\alpha x} + \frac{\tan^\alpha x}{1+\tan^\alpha x}\right\}dx = \frac{1}{2}\int_0^{\pi/2} dx = \frac{\pi}{4},$$

no matter the choice of $\alpha \in \mathbb{R}$. $\Diamond$

It is of course very difficult in a computer program to anticipate symmetry arguments like the one above, which is easier for the human mind. It is therefore recommended that the reader learns the principles of integration given in this chapter, and also tries to use them in exercises. But since the programs exist, one should of course also use them – to be familiar with that approach as well as to save time. When the computer programs give some very strange results, one has still the possibility of returning to the old tested methods.

Sometimes the interpretation of an integral as an area can give us a short cut to the solution.

Example 4.2 When $f(x) = \sqrt{1-x^2}$, $x \in [-1,1]$, the integral

$$\int_{-1}^{1} \sqrt{1-x^2}\,dx$$

is usually calculated by using the monotonous substitution $x = \sin t$, $t \in [-\pi/2, \pi/2]$. Then

$$\begin{aligned}\int_{-1}^{1} \sqrt{1-x^2}\,dx &= \int_{-\pi/2}^{\pi/2} \sqrt{1-\sin^2 t}\cdot\cos t\,dt = \int_{-\pi/2}^{\pi/2} |\cos t|\cdot\cos t\,dt = \int_{-\pi/2}^{\pi/2} \cos^2 t\,dt \\ &= \int_{-\pi/2}^{\pi/2} \frac{\cos 2t+1}{2}\,dt = \left[\frac{1}{4}\sin 2t + \frac{t}{2}\right]_{-\pi/2}^{\pi/2} = \frac{\pi}{2},\end{aligned}$$

with the important variant, using the symmetry,

$$\int_{-\pi/2}^{\pi/2} \cos^2 t\,dt = \int_{-\pi/2}^{\pi/2} \sin^2 t\,dt = \frac{1}{2}\int_{-\pi/2}^{\pi/2} \left\{\cos^2 t + \sin^2 t\right\}dt = \frac{1}{2}\int_{-\pi/2}^{\pi/2} dt = \frac{\pi}{2}.$$

A simpler method would be to notice that the graph of $f(x) = \sqrt{1-x^2}$, $x \in [-1,l]$, is a half circle in the upper half plane of radius 1. Hence the integral is

$$\int_{-1}^{1} \sqrt{1-x^2}\,dx = \pi\cdot 1^2 \cdot \frac{1}{2} = \frac{\pi}{2}. \qquad \Diamond$$

Finally, the symbol $\int f(x)\,dx$ is not unique. We can get another antiderivative by adding any arbitrary constant. It is, however, customary to leave out this arbitrary constant, as long as we are only trying to find one antiderivative. We shall therefore in this chapter always understand that a constant should be added in the end, and never write it. When it comes to the applications like in differential equations in Chapter 5 we must never forget this extra arbitrary constant.

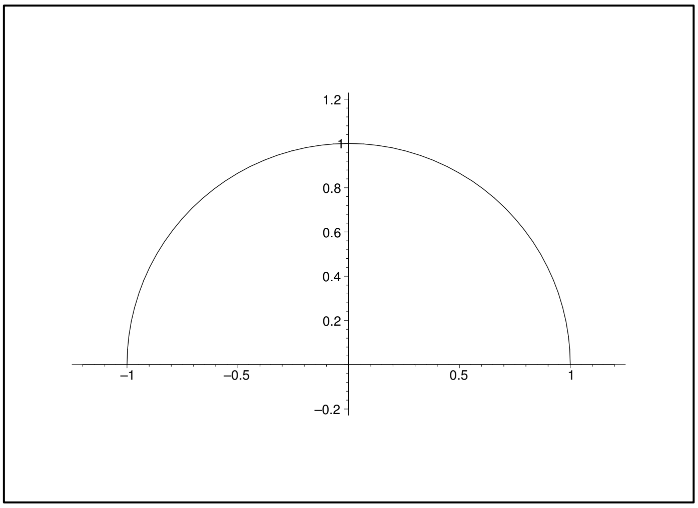

Figure 24: The graph of $f(x) = \sqrt{1-x^2}$, $x \in [-1,1]$.

Warning. Although experience shows that there are only few functions where one may be confused by this convention, students are nevertheless misled, when e.g. the antiderivative (typically) is not 0 for $x = 0$. For instance,

$$F(x) := \int e^x \, dx = e^x,$$

while we for the definite integral get

$$\int_0^x e^t \, dt = F(x) - F(0) = e^x - 1 \neq e^x.$$

This confusion occurs frequently in the students' exercises in various ways, so be careful here always to distinguish between the indefinite integral (or the antiderivative) and the definite integral.

4.2 A catalogue of standard antiderivatives.

During the last century many tables of antiderivatives have been worked out. We shall here limit ourselves to the most common ones.

Whenever possible we shall try to use the same structure as in the catalogue of the elementary derivatives. Although it is possible in many cases directly to reuse the results from Chapter 3, we are in other cases forced to derive a given new formula by using (79) and (80), so we cannot strictly maintain the same structure all over this section.

Power functions. If $f(x) = x^\alpha$, $x \in \mathbb{R}_+$, $\alpha \in \mathbb{R}$, then

$$(84)\quad F(x) = \int f(x)\, dx = \int x^\alpha dx = \begin{cases} \dfrac{1}{\alpha+1}\, x^{\alpha+1}, & \text{for } \alpha \neq -1, \\[2ex] \ln x, & \text{for } \alpha = -1. \end{cases}$$

Logarithm. If $f(x) = \ln x$, $x \in \mathbb{R}_+$, then

$$(85) \quad \int \ln x \, dx = x \cdot \ln x - x.$$

The easiest way to prove this formula is here to use (79) and (80),

$$\frac{d}{dx}\{x \cdot \ln x - x\} = 1 \cdot \ln x + x \cdot \frac{1}{x} - 1 = \ln x,$$

since we formally do not yet have the method of partial integration at hand.

The exponential functions. Apart for the arbitrary constant we obviously have

$$(86) \quad \int e^x dx = \int \exp x \, dx = e^x = \exp x.$$

Generally,

$$(87) \quad \int a^x dx = \int \exp(x \cdot \ln a)\, dx = \frac{1}{\ln a}\, a^x \qquad \text{for } a \in \mathbb{R}_+ \setminus \{1\},$$

which is proved by using (79) and (80), and

$$\int 1^x dx = \int dx = x \qquad \text{for } a = 1.$$

Trigonometric functions. It is well-known that

$$(88) \quad \begin{cases} \int \sin x \, dx = -\cos x, & x \in \mathbb{R}, \\[2mm] \int \cos x \, dx = \sin x, & x \in \mathbb{R}, \\[2mm] \dfrac{1}{\cos^2 x}\, dx = \int \left(1 + \tan^2 x\right)\, dx = \tan x, & x \neq \dfrac{\pi}{2} + p\pi, \quad p \in \mathbb{Z}, \\[2mm] \int \dfrac{1}{\sin^2 x}\, dx = \int \left(1 + \cot^2 x\right)\, dx = -\cot x, & x \neq p\pi, \qquad p \in \mathbb{Z}. \end{cases}$$

Furthermore,

$$(89) \quad \begin{cases} \int \tan x \, dx = \int \dfrac{\sin x}{\cos x}\, dx = -\ln|\cos x|, & x \neq \dfrac{\pi}{2} + p\pi, \quad p \in \mathbb{Z}, \\[2mm] \int \cot x \, dx = \int \dfrac{\cos x}{\sin x}\, dx = \ln|\sin x|, & x \neq p\pi, \qquad p \in \mathbb{Z}. \end{cases}$$

Hyperbolic functions. In just the same way as for the trigonometric functions we have

$$(90) \quad \begin{cases} \int \sinh x \, dx = \cosh x, & x \in \mathbb{R}, \\[2mm] \int \cosh x \, dx = \sinh x, & x \in \mathbb{R}, \\[2mm] \int \dfrac{1}{\cosh^2 x}\, dx = \int \left\{1 - \tanh^2 x\right\} dx = \tanh x, & x \in \mathbb{R}, \\[2mm] \int \dfrac{1}{\sinh^2 x}\, dx = \int \{\coth^x - 1\}\, dx = -\coth x, & x \in \mathbb{R} \setminus \{0\}. \end{cases}$$

Furthermore,

$$
(91) \quad \begin{cases} \int \tanh x \, dx = \int \dfrac{\sinh x}{\cosh x} \, dx = \ln \cosh x, & x \in \mathbb{R}, \\[2ex] \int \coth x \, dx = \int \dfrac{\cosh x}{\sinh x} \, dx = \ln |\sinh x|, & x \in \mathbb{R} \setminus \{0\}. \end{cases}
$$

Inverse trigonometric functions. It follows from Chapter 3,

$$
(92) \quad \begin{cases} \int \dfrac{dx}{\sqrt{1-x^2}} = \operatorname{Arcsin} x = -\operatorname{Arccos} x + c, & x \in]-1, 1[, \\[2ex] \int \dfrac{dx}{1+x^2} = \operatorname{Arctan} x = -\operatorname{Arccot} x + c, & x \in \mathbb{R}, \end{cases}
$$

where we exceptionally have added a constant c in order to emphasize that Arcsin x and $-\operatorname{Arccos} x$, resp. Arctan x and $-\operatorname{Arccot} x$, differ from each other by a constant. We proved in Chapter 3 that this constant was $c = \dfrac{\pi}{2}$ in both cases.

For completeness we mention that

$$\int \operatorname{Arcsin} x\,dx = x\cdot \operatorname{Arcsin} x + \sqrt{1-x^2}, \qquad x\in[-1,1],$$

$$\int \operatorname{Arccos} x\,dx = x\cdot \operatorname{Arccos} x - \sqrt{1-x^2}, \qquad x\in[-1,1],$$

$$\int \operatorname{Arctan} x\,dx = x\cdot \operatorname{Arctan} x - \frac{1}{2}\ln\left(1+x^2\right), \qquad x\in\mathbb{R},$$

$$\int \operatorname{Arccot} x\,dx = x\cdot \operatorname{Arccot} x + \frac{1}{2}\ln\left(1+x^2\right), \qquad x\in\mathbb{R},$$

but these results are also easily derived from the rule of partial integration, so it is no need to overburden one's memory with these formulæ.

Inverse hyperbolic functions. It follows from Chapter 3,

$$(93)\quad \begin{cases} \displaystyle\int \frac{dx}{\sqrt{x^2+1}} = \operatorname{Arsinh} x = \ln\left(x+\sqrt{x^2+1}\right), & x\in\mathbb{R},\\[2ex] \displaystyle\int \frac{dx}{\sqrt{x^2-1}} = \operatorname{Arcosh} x = \ln\left(x+\sqrt{x^2-1}\right), & x\in[1,+\infty[,\\[2ex] \displaystyle\int \frac{dx}{1-x^2} = \begin{cases} \operatorname{Artanh} x, & x\in]-1,1[,\\ \operatorname{Arcoth} x, & x\in\mathbb{R}\setminus[-1,1],\\ \displaystyle\frac{1}{2}\ln\left|\frac{x+1}{x-1}\right|, & x\in\mathbb{R}\setminus\{-1,1\}. \end{cases} \end{cases}$$

There is of course in the latter case a tendency in practice mostly to use

$$\int \frac{dx}{1-x^2} = \frac{1}{2}\ln\left|\frac{x+1}{x-1}\right|, \qquad x\neq -1,\ 1,$$

partly because ln is a more familiar function, and partly because its domain is the largest.

We mention for completeness only

$$\int \operatorname{Arsinh} x\,dx = x\cdot \operatorname{Arsinh} x - \sqrt{x^2+1}, \qquad x\in\mathbb{R},$$

$$\int \operatorname{Arcosh} x\,dx = x\cdot \operatorname{Arcosh} x + \sqrt{x^2+1}, \qquad x\in[1,+\infty[,$$

$$\int \operatorname{Artanh} x\,dx = x\cdot \operatorname{Artanh} x + \frac{1}{2}\ln\left(1-x^2\right), \qquad x\in]-1,1[,$$

$$\int \operatorname{Arcoth} x\,dx = x\cdot \operatorname{Arcoth} x - \frac{1}{2}\ln\left(x^2-1\right), \qquad x\notin[-1,1],$$

but these formulæ are always easily derived by using partial integration, so there is no need to remember them either.

Some special useful antiderivatives. Here we collect some important integrals shown previously,

$$(94)\quad \int \frac{dx}{1+x^2} = \begin{cases} \operatorname{Arctan} x, & x\in\mathbb{R},\\ -\operatorname{Arccot} x, & x\in\mathbb{R}, \end{cases}$$

and

(95) $$\int \frac{dx}{\sqrt{1-x^2}} = \begin{cases} \text{Arcsin}\, x, & x \in [-1,1], \\ -\text{Arccos}\, x, & x \in [-1,1], \end{cases}$$

(96) $$\int \frac{dx}{\sqrt{x^2-1}} = \text{Arcosh}\, x = \ln\left(x + \sqrt{x^2-1}\right), \qquad x \in [1,+\infty[,$$

(97) $$\int \frac{dx}{\sqrt{x^2+1}} = \text{Arsinh}\, x = \ln\left(x + \sqrt{x^2+1}\right), \qquad x \in \mathbb{R}.$$

They will be necessary in more advanced integration problems.

4.3 Simple rules of integration.

We shall as usual not bother with the arbitrary constants, which should be added whenever one sees the symbol $\int \cdots dx$.

Linearity. Let a, $b \in \mathbb{R}$ be constants and f, g integrable functions, defined in the same interval. Then

(98) $$\int \{a\, f(x) + b\, g(x)\}\, dx = a \int f(x)\, dx + b \int g(x)\, dx.$$

Partial integration, or integration by parts. Let f and g be two functions defined in the same interval and with their respective antiderivatives F and G (any fixed of them), i.e.

$$F(x) = \int f(x)\, dx \qquad \text{and} \qquad G(x) = \int g(x)\, dx.$$

When the product $F(x) \cdot G(x)$ is differentiated we get

$$\begin{aligned} \frac{d}{dx}\{F(x)G(x)\} &= F'(x) \cdot G(x) + F(x) \cdot G'(x) \\ &= f(x) \cdot G(x) + F(x) \cdot g(x). \end{aligned}$$

By an integration, followed by a rearrangement we obtain the *rule of integration by parts*

(99) $$\int f(x) \cdot G(x)\, dx = F(x)G(x) - \int F(x)g(x)\, dx.$$

This rule is extremely useful, both in simple and in more advanced problems. In the simple problems we use it typically to differentiate "bad" factors in the integrand like the logarithm or inverse trigonometric or inverse hyperbolic functions, just to mention a few of the applications. We show one of these applications in the following example.

Example 4.3 When we choose

$$f(x) = 1 \qquad \text{and} \qquad G(x) = \text{Arcsin}\, x, \qquad x \in]-1,1[,$$

we see that

$$F(x) = x \qquad \text{and} \qquad g(x) = \frac{1}{\sqrt{1-x^2}}, \qquad x\,]-1,1[.$$

Inserting these in (99) we get

$$\begin{aligned}\int \operatorname{Arcsin} x \, dx &= \int 1 \cdot \operatorname{Arcsin} x \, dx = \int f(x)G(x)\, dx \\ &= F(x)G(x) - \int F(x)g(x)\, dx = x \cdot \operatorname{Arcsin} x - \int x \cdot \frac{1}{\sqrt{1-x^2}}\, dx.\end{aligned}$$

When we continue by calculating the latter integral we finally get

$$\int \operatorname{Arcsin} x \, dx = x \cdot \operatorname{Arcsin} x + \sqrt{1-x^2}. \qquad \Diamond$$

The rule of integrating by parts is far more important than the example above shows. In practice most integrals cannot be given an exact expression, but by using integration by parts we may be able to derive some reasonable approximations. We shall here illustrate the idea by proving a variant of *Taylor's formula.*

Example 4.4 Let $f \in C^\infty$ be an infinitely often differentiable function in a neighbourhood of a point $a \in \mathbb{R}$. (Once we have been through the following argument, it is easy to relax the assumption to $f \in C^n$ for some $n \in \mathbb{N}$.) Then we have in the same neighbourhood,

$$f(x) = f(a) + \int_a^x f'(t)\, dt = f(a) + \int_a^x 1 \cdot f'(t)\, dt.$$

Since we integrate with respect to t, we consider x as a *constant* in the formula above. The *dirty trick* here is therefore that one particular antiderivative of 1 with respect to t is

$$\int 1\, dt = t - x.$$

(Check this formula by differentiating with respect to t, where x is considered as a constant, and we see that the formula is correct!)

When we integrate by parts (the constant 1 is integrated, while $f'(t)$ is differentiated) we get

$$\begin{aligned}f(x) &= f(a) + \int_a^x f'(t)\, dt = f(a) + \int_a^x 1 \cdot f'(t)\, dt \\ &= f(a) + [(t-x)f'(t)]_{t=a}^x - \int_a^x (t-x)f''(t)\, dt \\ &= f(a) + (x-a)f'(a) - \int_a^x (t-x)f''(t)\, dt \\ &= f(a) + f'(a)\cdot(x-a) + \int_x^a (t-x)f''(t)\, dt.\end{aligned}$$

When we repeat this integration by parts on the latter integral we get

$$\begin{aligned}\int_x^a (t-x)f''(t)\, dt &= \left[\frac{1}{2}(t-x)^2 f''(t)\right]_x^a - \frac{1}{2}\int_x^a (t-x)^2 f^{(3)}(t)\, dt \\ &= \frac{1}{2} f''(a)\cdot(x-a)^2 - \frac{1}{2}\int_x^a (t-x)^2 f^{(3)}(t)\, dt.\end{aligned}$$

Repeating this method we get in the next step in the given neighbourhood,

$$-\frac{1}{2}\int_x^a (t-a)f^{(3)}(t)\, dt = -\frac{1}{2\cdot 3}\left[(t-x)^3 f^{(3)}(t)\right]_{t=x}^a + \frac{1}{2\cdot 3}\int_x^a (t-x)^3 f^{(4)}(t)\, dt.$$

By insertion we get in the given neighbourhood,

(100) $$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \frac{1}{3!}\int_x^a (t-x)^3 f^{(4)}(t)\,dt.$$

The idea of a *Taylor's formula* like e.g. (100) is that if we put $|x-a| = \varepsilon$, then we have the following estimate of the remainder term

$$\left|\frac{1}{3!}\int_x^a (t-x)^3 f^{(4)}(t)\,dt\right| \le \frac{1}{3!}\sup_{|t-a|\le\varepsilon}\left\{\left|f^{(4)}(t)\right|\cdot|x-a|^4\right\} = \frac{1}{3!}\,\varepsilon^4 \sup_{|t-a|\le\varepsilon}\left|f^{(4)}(t)\right|.$$

Since $f^{(4)}(t)$ is continuous, this error can be made as small as we want by choosing a sufficiently small $\varepsilon > 0$.

It is obvious how we shall proceed, as long as $f^{(n)}(t)$ occurring in the remainder term is defined. We shall return to this idea in Chapter 6.

We see that we can derive a variant of *Taylor's formula* without using *Rolle's theorem* or the *mean value theorem*.

Example 4.5 As a third application we want to approximate the integral

$$F(x) = \int_0^x \exp\left(-t^2\right) dt, \qquad \text{for } x \in [0,1].$$

When we successively integrate by parts – always differentiating the exponential function and integrating the monomial – we get

$$\begin{aligned} F(x) &= \int_0^x 1 \cdot e^{-t^2} dt = \left[t \cdot e^{-t^2}\right]_0^x + 2\int_0^x t^2 e^{-t^2} dt = x \cdot e^{-x^2} + \frac{2}{3}\left[t^3 e^{-t^2}\right]_0^x + \frac{4}{3}\int_0^x t^4 e^{-t^2} dt \\ &= x \cdot e^{-x^2} + \frac{2}{3}\, x^3 e^{-x^2} + \frac{4}{3 \cdot 5}\left[t^5 e^{-t^2}\right]_0^x + \frac{8}{3 \cdot 5}\int_0^x t^6 e^{-t^2} dt \\ &= e^{-x^2}\left\{x + \frac{2}{1 \cdot 3}\, x^3 + \frac{2^2}{1 \cdot 3 \cdot 5}\, x^5\right\} + \frac{2^3}{1 \cdot 3 \cdot 5}\int_0^x t^6 e^{-t^2} dt. \end{aligned}$$

Proceeding in this way we get

$$\begin{aligned} F(x) &= e^{-x^2}\left\{x + \frac{2}{1 \cdot 3}\, x^3 + \frac{2^2}{1 \cdot 3 \cdot 5}\, x^5 + \cdots + \frac{2^n}{1 \cdot 3 \cdot 5 \cdots (2n+1)}\, x^{2n+1}\right\} \\ &\qquad + \frac{2^{n+1}}{1 \cdot 3 \cdot 5 \cdots (2n+1)}\int_0^x t^{2n+2} e^{-t^2} dt. \end{aligned}$$

The factor in front of the integral tends to 0 for $n \to +\infty$,

$$\frac{2^{n+1}}{1 \cdot 3 \cdot 5 \cdots (2n+1)} = \frac{2}{1} \cdot \frac{2}{3} \cdot \frac{2}{5} \cdot \frac{2}{2n+1} = \frac{2^{2n+1} n!}{(2n+1)!},$$

and since we have assumed that $x \in [0,1]$, the integral is trivially < 1. Hence,

$$F(x) \approx e^{-x^2}\left\{x + \frac{2}{1 \cdot 3}\, x^3 + \frac{2^2}{1 \cdot 3 \cdot 5}\, x^5 + \cdots + \frac{2^n}{1 \cdot 3 \cdot 5 \cdots (2n+1)}\, x^{2n+1}\right\} \qquad \text{for } x \in [0,1].$$

By estimating the factor in front of the integral we also get an estimate of the error made by this approximation.

We have above given three different applications of the *rule of partial integration* in order to emphasize its great importance.

4.4 Integration by substitution

There are two possibilities, one where the substitution is more or less given by the structure of the integrand, and one where one has to introduce the substitution oneself.

First rule of integration by substitution. Let $f \in C^2$ and $g \in C^0$. Then

$$(101) \quad \int g(f(x)) \cdot f'(x)\, dx = \text{“}\int g(f(x))\, df(x)\text{”} = \int_{u=f(x)} g(u)\, du.$$

Proof. The rule is proved by differentiation,

$$\begin{aligned} \frac{d}{dx}\left\{\int_{u=f(x)} g(u)\, du\right\} &= \frac{d}{du}\left\{\int g(u)\, du\right\} \cdot \left.\frac{du}{dx}\right|_{u=f(x)} = g(u)|_{u=f(x)} \cdot f'(x) \\ &= g(f(x)) \cdot f'(x). \qquad \Box \end{aligned}$$

Example 4.6 Let $n \in \mathbb{N}_0$. A typical application is

$$\int \sin^n x \cdot \cos x \, dx = \text{``}\int \sin^n x \, d \sin x\text{''} = \frac{1}{n+1} \sin^{n+1} x.$$

Another one is

$$\int \tan^n x \cdot \left\{1 + \tan^2 x\right\} dx = \text{``}\int \tan^n x \, d \tan x\text{''} = \frac{1}{n+1} \tan^{n+1} x.$$

By a rearrangement of the latter formula we get the recursion formula

$$(102) \quad \int \tan^{n+2} x \, dx = \frac{1}{n+1} \tan^{n+1} x - \int \tan^n x \, dx,$$

hence when n is replaced by $n-2$ in (102),

$$\int \tan^n x \, dx = \frac{1}{n-1} \tan^{n-1} x - \int \tan^{n-2} x \, dx \qquad \text{for } n \neq 1.$$

Since

$$\int \tan^1 x \, dx = -\ln|\cos x| \qquad \text{and} \qquad \int \tan^0 x \, dx = \int 1 \, dx = x,$$

it is easy to find an explicit expression for $\int \tan^n x \, dx$ by considering n either even or odd. $\lozenge$

Second rule of integration by substitution. Let $x = f(t)$ be a *strictly monotonous* C^1-function. Then $f(t)$ has an inverse $t = f^{-1}(x)$, and

$$(103) \quad \int g(x) \, dx = \int_{t=f^{-1}(x)} g(f(t)) \, f'(t) \, dt.$$

When we compare the two rules we see that they are almost identical. The difference is that we in (101) use that $f'(x)$ already exists as a factor in the integrand, so it is of no importance whether $u = f(x)$ is monotonous or not. It is on the other hand obvious from (103) that the inverse function $t = f^{-1}(x)$ in this case *must* exist. Therefore we *must* in (103) assume that the chosen substitution $t = f(x)$ indeed is strictly monotonous and of class C^1.

When we apply this rule, the general principle is that if some "ugly term" occurs in the integrand, then try to use it as a substitution. It does not work in all cases, but there is a reasonable chance that it does. In other cases we may look for some functional equation which resembles the integrand, if x is substituted by a function which enters the functional equation.

Example 4.7 Assuming that $x > 0$ we shall find an exact expression for the antiderivative

$$\int \frac{1}{x^3} \exp\left(\frac{1}{x}\right) dx.$$

The function $\exp\left(\frac{1}{x}\right)$ is far from nice, but fortunately $x = f(t) := \frac{1}{t}$ is monotonous for $x > 0$ and $t > 0$, where

$$dx = -\frac{1}{t^2} \, dt \qquad \text{and} \qquad t = f^{-1}(x) = \frac{1}{x}.$$

Thus we get by using this substitution,

$$\begin{aligned}\int \frac{1}{x^3}\exp\left(\frac{1}{x}\right)dx &= \int_{t=1/x} t^3\exp(t)\cdot\left(-\frac{1}{t^2}\right)dt = -\int_{t=1/x} t\,e^t dt = -\left[t\cdot e^t - e^t\right]_{t=1/x}\\ &= \exp\left(\frac{1}{x}\right)\cdot\left\{1-\frac{1}{x}\right\} = \frac{x-1}{x}\cdot\exp\left(\frac{1}{x}\right). \qquad \Diamond\end{aligned}$$

The second rule of integration by substitution is in particular chosen when we have a square root in the integrand. We mention here four very typical substitutions.

1) If the integrand contains $\sqrt{x-a}$ for $x>a$, we put

$$t = f(x) = \sqrt{x-a} > 0,$$

so

$$x = f^{-1}(t) = t^2 + a \qquad \text{and} \qquad dx = 2t\,dt.$$

2) If the integrand contains $\sqrt{1-x^2}$ for $-1<x<1$, we choose

$$x = f(t) = \sin t \qquad \text{for } -\frac{\pi}{2} < t < \frac{\pi}{2}.$$

Then

$$dx = \cos t\,dt \qquad \text{and} \qquad \sqrt{1-x^2} = +\cos t,$$

and

$$t = f^{-1}(x) = \operatorname{Arcsin} x, \qquad -1<x<1.$$

Here it is also possible to choose the substitution $x=\cos t$ for $0<t<\pi$. The details are left to the reader.

3) If the integrand contains $\sqrt{1+x^2}$, $x\in\mathbb{R}$, we choose the substitution

$$x = f(t) = \sinh t, \qquad \text{because then } \sqrt{1+x^2} = \cosh t.$$

Here we use that $\cosh^2 t - \sinh^2 t = 1$ and $\cosh t \geq 0$. It follows that

$$dx = \cosh t\,dt \qquad \text{and} \qquad t = f^{-1}(x) = \operatorname{Arsinh} x = \ln\left(x+\sqrt{x^2+1}\right).$$

4) If the integrand contains $\sqrt{x^2-1}$ for $x>1$ (or for $x<-1$ with similar calculations), we choose the substitution

$$x = f(t) = \cosh t, \quad t>0, \qquad \text{because then } \sqrt{x^2-1} = \sinh t.$$

It follows that

$$dx = \sinh t\,dt \qquad \text{and} \qquad t = f^{-1}(x) = \operatorname{Arcosh} x = \ln\left(x+\sqrt{x^2-1}\right).$$

Before the age of advanced pocket calculators and computer programs the student was supposed to master quite a lot of more special cases of integrals than those already mentioned above. I think that it suffices with the cases mentioned in this chapter, supplied with what one can get from computer programs.

4.5 Complex decomposition of fractions of polynomials.

A typical integration problem is given when the integrand is a fraction between two polynomials. The standard procedure is first to decompose the fraction into a sum of special fractions, which can be integrated. Such a decomposition has independent interest, so I include here a section on this subject alone, and then applies it in the next section, where we integrate the fraction.

Let $\dfrac{P(x)}{Q(x)}$ be a fraction between two polynomials with no common complex roots. We shall decompose this fraction into a sum of fractions, each of which can easily be integrated.

Procedure.

1) If the degree of the numerator is $\geq$ the degree of the denominator, one starts with a division,

$$\frac{P(x)}{Q(x)} = P_1(x) + \frac{R(x)}{Q(x)},$$

where $P_1(x)$ and $R(x)$ are polynomials, and the later polynomial $R(x)$ has degree $>$ the degree of the denominator $Q(x)$.

We save the polynomial $P_1(x)$ until the final step and continue only with the fraction $\dfrac{R(x)}{Q(x)}$.

2) Find all the complex roots and their multiplicity of the denominator $Q(x)$, such that we have

$$Q(x) = c \cdot (x-a_1)^{p_1} \cdots (x-a_k)^{p_k}\,.$$

Check that the sum $p_1 + \cdots + p_k$ of all exponents is equal to the degree of $Q(x)$.
If furthermore $Q(x)$ is real, check that complex conjugated roots are of the same multiplicity.

3) Choose the simplest of the two polynomials $P(x)$ and $R(x)$. The following method will give the same result, no matter which one of these is chosen. Since it is theoretically most correct to write $R(x)$ here, we shall do it in the rest of the procedure, although $P(x)$ would give the same results. The important thing here is that we above have transferred the polynomial $P_1(x)$ to the final step. (Otherwise we would in the end get a wrong result.)

4) The fraction can now be written

$$(104)\quad \frac{R(x)}{Q(x)} = \frac{1}{c}\cdot\frac{R(x)}{(x-a_1)^{p_1}\cdots(x-a_k)^{p_k}}.$$

The coefficient of the fraction

$$\frac{1}{(x-a_1)^{p_1}}$$

is found by "putting your hand over" the factor $(x-a_1)^{p_1}$ in the denominator in (104) (i.e. technically we are multiplying (104) by $(x-a_1)^{p_1}$) and then let $x=a_1$ in the remainder, i.e.

$$b_{1,p_1} = \frac{1}{c}\cdot\left.\frac{R(x)}{(x-a_2)^{p_2}\cdots(x-a_k)^{p_k}}\right|_{x=a_1}.$$

The resulting fraction

$$\frac{b_{1,p_1}}{(x-a_1)^{p_1}}$$

is saved for the final result.

5) Repeat step 4 on each of the other factors

$$(x-a_2)^{p_2}\,,\ \dots,\ (x-a_k)^{p_k}\,,$$

in the denominator and save the results for the final step.

6) All the fractions found in step 4 and step 5 are now subtracted from $\dfrac{R(x)}{Q(x)}$ followed by a reduction,

$$\frac{1}{c}\cdot\frac{R(x)}{(x-a_1)^{p_1}\cdots(x-a_k)^{p_k}} - \frac{b_{1,p_1}}{(x-a_1)^{p_1}} - \cdots - \frac{b_{k,p_k}}{(x-a_1)^{p_k}} = \frac{1}{d}\cdot\frac{R_1(x)}{(x-a_1)^{q_1}\cdots(x-a_k)^{q_k}}.$$

Here we must have

$$q_1 < p_1,\ \dots,\ q_k < p_k,$$

so these inequalities should be checked.

7) Repeat the steps 4, 5 and 6 on the reduced fraction

$$\frac{1}{d} \cdot \frac{R_1(x)}{(x-a_1)^{q_1} \cdots (x-a_k)^{q_k}}.$$

In each step the results should be transferred to the final step. Since $q_1 + \cdots + q_k < p_1 + \cdots + p_k$, this process must necessarily stop after a finite number of iterations.

8) Collect all the found fractions and the polynomial $P_1(x)$ from step 1. The result is the wanted decomposition.

I have above described the mechanical procedure, but there will often exist shortcuts, which are faster. These are, however, difficult to describe for a computer.

Example 4.8 *Find the decomposition of the fraction* $\dfrac{1}{x^4-1}$.

1) *Standard procedure* as described above.
The denominator has the simple roots 1, i, -1, -i, hence

$$\begin{aligned}
\frac{1}{x^4-1} &= \frac{1}{(x-1)(x-\mathrm{i})(x+1)(x+\mathrm{i})} \\
&= \frac{1}{(1-\mathrm{i})(1+1)(1+\mathrm{i})} \cdot \frac{1}{x-1} + \frac{1}{(\mathrm{i}-1)(\mathrm{i}+1)(\mathrm{i}+\mathrm{i})} \cdot \frac{1}{x-\mathrm{i}} \\
&\quad + \frac{1}{(-1-1)(-1-\mathrm{i})(-1+\mathrm{i})} \cdot \frac{1}{x+1} + \frac{1}{(-\mathrm{i}-1)(-\mathrm{i}-\mathrm{i})(-\mathrm{i}+1)} \cdot \frac{1}{x+\mathrm{i}} \\
&= \frac{1}{4} \cdot \frac{1}{x-1} - \frac{1}{4\mathrm{i}} \cdot \frac{1}{x-\mathrm{i}} - \frac{1}{4} \cdot \frac{1}{x+1} + \frac{1}{4\mathrm{i}} \cdot \frac{1}{x+\mathrm{i}} \\
&= \frac{1}{4} \cdot \frac{1}{x-1} - \frac{1}{4} \cdot \frac{1}{x+1} - \frac{1}{4\mathrm{i}} \left\{ \frac{1}{x-\mathrm{i}} - \frac{1}{x+\mathrm{i}} \right\} \\
&= \frac{1}{4} \cdot \frac{1}{x-1} - \frac{1}{4} \cdot \frac{1}{x+1} - \frac{1}{2} \cdot \frac{1}{x^2+1}.
\end{aligned}$$

2) *Alternatively* we first notice that

$$x^4 - 1 = \left(x^2\right)^2 - 1 = \left(x^2+1\right)\left(x^2-1\right).$$

By first writing $u = x^2$, then decomposing after u, and finally after x we get the simpler calculations

$$\begin{aligned}
\frac{1}{x^4-1} &= \frac{1}{u^2-1} = \frac{1}{(u-1)(u+1)} = \frac{1}{2}\,\frac{1}{u-1} - \frac{1}{2}\,\frac{1}{u+1} = \frac{1}{2}\,\frac{1}{x^2-1} - \frac{1}{2}\,\frac{1}{x^2+1} \\
&= \frac{1}{2}\,\frac{1}{(x-1)(x+1)} - \frac{1}{2}\,\frac{1}{x^2+1} = \frac{1}{4}\,\frac{1}{x-1} - \frac{1}{4}\,\frac{1}{x+1} - \frac{1}{2}\,\frac{1}{x^2+1}. \qquad \Diamond
\end{aligned}$$

4.6 Integration of a fraction of two polynomials.

Let $P(x)$ and $Q(x)$ be real polynomials. We shall describe a procedure of finding the integral $\dfrac{P(x)}{Q(x)}\,dx$.

Remark 4.1 Occasionally the choices of $P(x)$ and $Q(x)$ will cause that the pocket calculators TI-92 and TI-89 give some very strange results. Therefore it is recommended to learn the method of this section. $\Diamond$

Procedure.

1) Decompose the fraction $\dfrac{P(x)}{Q(x)}$ as described in Section 4.5. This means that $\dfrac{P(x)}{Q(x)}$ is written as a sum of a polynomial and fractions of the type $\dfrac{c}{(x-a)^p}$, where the root of the denominator $a \in \mathbb{C}$ may be complex as well as the constant $c \in \mathbb{C}$.

2) The polynomial $P_1(x)$ is integrated as usual.

3) Fractions $\dfrac{c}{(x-a)^p}$, where $p > 1$, are integrated in the usual way, no matter whether a is real or complex,

$$\int \frac{c}{(x-a)^p}\,dx = -\frac{c}{p-1}\cdot\frac{1}{(x-a)^{p-1}}.$$

If $P(s)$ and $Q(x)$ are real, and a is complex in $\dfrac{c}{(x-a)^p}$, then the decomposition of $\dfrac{P(x)}{Q(x)}$ also contains the complex conjugated fraction $\dfrac{\overline{c}}{(x-\overline{a})^p}$. By integration, followed by a reduction we get

$$\begin{aligned}\int\left\{\frac{c}{(x-a)^p}+\frac{\overline{c}}{(x-\overline{a})^p}\right\}dx &= -\frac{1}{p-1}\left\{\frac{c}{(x-a)^{p-1}}+\frac{\overline{c}}{(x-\overline{a})^{p-1}}\right\}\\ &= -\frac{2}{p-1}\,\mathrm{Re}\left\{\frac{c}{(x-a)^{p-1}}\cdot\frac{(x-\overline{a})^{p-1}}{(x-\overline{a})^{p-1}}\right\}\\ &= -\frac{2}{p-1}\cdot\frac{\mathrm{Re}\left\{c\cdot(x-\overline{a})^{p-1}\right\}}{\{x^2-2\,\mathrm{Re}\,a\cdot x+|a|^2\}^{p-1}}.\end{aligned}$$

4) If $p = 1$, and $a \in \mathbb{R}$ is real, then

$$\int \frac{c}{x-a}\,dx = c\cdot\ln|x-a|.$$

5) If $p = 1$ and $a = a_1 + \mathrm{i}\,a_2 \in \mathbb{C}$ is complex, then

$$\frac{c}{x-a} = \frac{c}{(x-a_1)-\mathrm{i}\,a_2}\cdot\frac{x-a_1+\mathrm{i}\,a_2}{(x-a_1)+\mathrm{i}\,a_2} = \frac{c\,(x-a_1)+\mathrm{i}\,ca_2}{(x-a_1)^2+a_2^2},$$

so

$$\begin{aligned}\int\frac{c}{x-a}\,dx &= c\int\frac{x-a_1}{(x-a_1)^2+a_2^2}\,dx+\mathrm{i}\,c\int\frac{1}{\left(\dfrac{x-a_1}{a_2}\right)^2+1}\cdot\frac{1}{a_2}\,dx\\ &= \frac{c}{2}\ln\left\{(x-a_1)^2+a_2^2\right\}+\mathrm{i}\,c\,\mathrm{Arctan}\left(\frac{x-a_1}{a_2}\right).\end{aligned}$$

6) The final result is obtained by adding all the results from the steps 2, 3, 4 and 5.

Example 4.9 We found in Section 4.5 the decomposition

$$\frac{1}{x^4-1} = \frac{1}{4}\,\frac{1}{x-1} - \frac{1}{4}\,\frac{1}{x+1} - \frac{1}{2}\,\frac{1}{x^2+1}, \qquad x \neq \pm 1.$$

Hence by integration,

$$\int \frac{1}{x^4-1}\,dx = \frac{1}{4}\,\ln\left|\frac{x-1}{x+1}\right| - \frac{1}{2}\operatorname{Arctan} x, \qquad x \neq \pm 1, \qquad \Diamond$$

The method using the complex decomposition is not necessary, if one is only interested in the antiderivative of $\dfrac{P(x)}{Q(x)}$. We shall illustrate this alternative method by an example, in which it is seen that the most important thing is only to find the roots and their multiplicity of the denominator.

Example 4.10 *Find the antiderivative*

$$\int \frac{ax^2+bx+c}{\left(x^2+1+1\right)(x+1)}\,dx, \qquad \textit{for } x>-1,$$

where a, b and c are given constants.

If we performed a decomposition of the integrand, we would get some linear expression in the three terms

$$\frac{1}{x+1}, \qquad \frac{x}{x^2+1}, \qquad \frac{1}{x^2+1}.$$

There must therefore exist new constants α, β, γ, such that

$$(105) \quad \int \frac{ax^2+bx+c}{\left(x^2+1\right)(x+1)}\,dx = \alpha\,\ln(x+1)+\beta\,\ln\left(x^2+1\right)+\gamma\,\mathrm{Arctan}\,x.$$

We shall only find the constants α, β, γ.

When we differentiate (105) we get

$$\frac{ax^2+bx+c}{\left(x^2+1\right)(x+1)} = \frac{\alpha}{x+1}+\frac{2\beta\,x}{x^2+1}+\frac{\gamma}{x^2+1},$$

hence,

$$\begin{aligned} ax^2+bx+c &= \alpha\left(x^2+1\right)+2\beta x(x+1)+\gamma(x+1) \\ &= (\alpha+2\beta)x^2+(2\beta+\gamma)x+(\alpha+\gamma). \end{aligned}$$

By identification of the coefficients we get

$$\alpha+2\beta=a, \qquad 2\beta+\gamma=b \quad \text{and} \quad \alpha+\gamma=c,$$

hence

$$\alpha=\frac{a-b+c}{2}, \qquad \beta=\frac{a+b-c}{4} \quad \text{and} \quad \gamma=\frac{-a+b+c}{2}. \qquad \Diamond$$

4.7 Integration of trigonometric polynomials.

Let m, $n \in \mathbb{N}_0$. We shall here describe a method of calculating integrals of the type

$$\int \sin^m x \cdot \cos^n x\,dx,$$

where we say that the trigonometric polynomial $\sin^m x \cdot \cos^n x$ (the integrand) is of degree $m+n$.

There are two cases, namely *odd* or *even* degree of the integrand. These cases are again divided into two subcases. Hence,

1) $m+n$ odd.

   a) m even and n odd, i.e. $m=2p$ and $n=2q+1$, where p and $q \in \mathbb{N}_0$.

   b) m odd and n even, i.e. $m=2p+1$ and $n=2q$, where p and $q \in \mathbb{N}_0$.

2) $m+n$ even.

a) m and n are both odd, i.e. $m = 2p+1$ and $n = 2q+1$, where p and $q \in \mathbb{N}_0$.

b) m and n are both even, i.e. $m = 2p$ and $n = 2q$, where p and $q \in \mathbb{N}_0$.

The most difficult case is 2b), where both m and n are even.

Procedure.

1) a) $m = 2p$ and $n = 2p+1$.
Use the substitueion $u = \sin x$ (first rule) corresponding to $m = 2p$ *even*,

$$\begin{aligned}\int \sin^{2p} x \cdot \cos^{2q+1} x\, dx &= \int \sin^{2p} x \cdot \cos^{2q} x \cdot \cos x\, dx \\ &= \int \sin^{2p} x \cdot \left(1-\sin^2 x\right)^q d\sin x \\ &= \int_{u=\sin x} u^{2p}\left(1-u^2\right)^q du,\end{aligned}$$

where it is easy to integrate the polynomial in u, once p and q are given.

b) $m = 2p+1$ and $n = 2q$.
Use the substitution $u = \cos x$ (first rule) corresponding to $n = 2q$ *even*,

$$\begin{aligned}\int \sin^{2p+1} x \cdot \cos^{2q} x\, dx &= \int \sin^{2p} x \cdot \cos^{2q} x \cdot \sin x\, dx \\ &= \int \left(1-\cos^2 x\right)^p \cos^{2q} x \cdot (-1)\, d\cos x \\ &= -\int_{u=\cos x} \left(1-u^2\right)^p u^{2q} du,\end{aligned}$$

where it is easy to integrate the polynomial in u, once p and q are given.

2) When $m+n$ is *even*, the trick is to use the double angle $2x$, whence the degree is halved. We shall use the well-known formulæ from high school,

$$\cos^2 = \frac{1}{2}(1+\cos 2x), \qquad \sin^2 x = \frac{1}{2}(1-\cos 2x), \qquad \sin x \cdot \cos x = \frac{1}{2}\sin 2x,$$

when we reduce the integrand.

a) $m = 2p+1$ and $n = 2q+1$ are both *odd*.
The integrand is rewritten in the following way

$$\begin{aligned}\sin^{2p+1} x \cdot \cos^{2q+1} x &= \left(\sin^2 x\right)^p \cdot \left(\cos^2 x\right)^q \cdot \sin x \cdot \cos x \\ &= \left\{\frac{1}{2}(1-\cos 2x)\right\}^p \left\{\frac{1}{2}(1+\cos 2x)\right\}^q \cdot \frac{1}{2}\sin 2x.\end{aligned}$$

By insertion we reduce the problem to a special case of 1b). By the substitution $u = \cos 2x$ (first rule) we finally get

$$\begin{aligned}\int \sin^{2p+1} x \cdot \cos^{2q+1} x\, dx &= \frac{1}{2^{p+q+1}} \int (1-\cos 2x)^p (1+\cos 2x)^q \sin 2x\, dx \\ &= \frac{1}{2^{p+q+1}} \int (1-\cos 2x)^p (1+\cos 2x)^q \cdot \left(-\frac{1}{2}\right) d\cos 2x \\ &= -\frac{1}{2^{p+q+2}} \int_{u=\cos 2x} (1-u)^p (1+u)^q\, du,\end{aligned}$$

so the problem is again reduced to the integration of a polynomial.

b) $m = 2p$ and $n = 2q$ are both *even*.
In this case there is no solution formula like in the other three cases, but we still have a procedure, which reduces the problem to a finite number of problems of the types 1a) or 2b) of smaller degree. The final result is obtained after a finite number of steps.

The clue is to rewrite the integrand in the following way,

$$(106)\quad \sin^{2p} x \cdot \cos^{2q} x = \left\{\frac{1}{2}\,(1 - \cos 2x)\right\}^p \left\{\frac{1}{2}\,(1 + \cos 2x)\right\}^q .$$

The left hand side of (106) is of degree $2p + 2q$ in x.
The right hand side of (106) is a trigonometric polynomial of the smaller degree $p + q$ in $2x$.

When the right hand side of (106) has been calculated, each term must be handled separately, depending on whether the degree $j\ (\leq p + q)$ is *odd*, in which case we are in case 1a), or *even*, in which case we are in the worst case 2b), though of smaller degree.

5 Simple Differential Equations

5.1 Introduction.

One of the most rewarding applied topics in mathematics is the theory of differential equations. These are dominating in almost every physical model. For the same reason they have been studied for more than a century.

In earlier times one demanded that the students knew a lot about differential equations, including nonlinear differential equations of autonomous type, of Bernoulli type, and of Riccati type. These equations are no more common in the general teaching of Calculus. The only nonlinear differential equations which still are treated in the elementary courses are the ones in which the variables can be separated (Section 5.2). Any other differential equation here is linear.

We shall here treat the following types of linear differential equations:

Section 5.3. Linear differential equations of first order and variable coefficients.

Section 5.4. Linear differential equations of order n and constant coefficients.

Section 5.5. Euler's differential equation.

Section 5.6. Linear differential equations of second order and variable coefficients.

Section 5.6 is a little complicated, and one cannot expect it to be included in every first course in Calculus.

We focus on the linear differential equations for two reasons. First they occur extremely often in physical models, in particular differential equations of first and second order, and secondly they can in principle be solved by just as many successive integrations as the order of the differential equation. Notice the phrase *"in principle"*, because it is not always easy to show how one gets a solution formula. The main point here is that *linear* differential equations can be considered as variants of the problem of integration

$$\frac{df}{dx} = g(x) \qquad \text{with the solution} \qquad f(x) = \int g(x)\,dx + c,$$

where we unlike in Chapter 4 from now on *always* explicitly add the arbitrary constant c.

Hence integration is an indispensable prerequisite for linear differential equations.

We shall first consider the nonlinear differential equations of first order, which can be solved by separation of the variables.

5.2 Differential equations which can be solved by separation.

Let $f(x) \neq 0$ and $g(t)$ be continuous functions in different variables. We shall consider differential equations of the type

$$(107)\quad \frac{dx}{dt} = \frac{g(t)}{f(x)}.$$

Since $f(x) \neq 0$, this equation is equivalent to

$$(108)\quad g(t) = f(x)\,\frac{dx}{dt} = \frac{d}{dt}\left\{\int f(x)\,dx\right\},$$

where we implicitly have used the first rule of substitution, assuming as we may that $x = x(t)$ is a function i t.

When (108) is integrated with respect to t we get with an arbitrary constant c that the complete solution of (107) is implicitly given by

$$(109)\quad F(x) = \int f(x)\,dx = \int g(t)\,dt + c = G(t) + c.$$

The nonlinear structure is seen by the fact that we afterwards get the trouble of finding x explicitly as a function of t by using (109), i.e. formally

$$(110)\quad x = F^{-1}(G(t) + c), \qquad c \in \mathbb{R} \text{ an arbitrary constant,}$$

whenever the inverse F^{-1} of F exists. This is in general not an easy task. We shall briefly sketch an approximation method in Chapter 6 of this problem.

When (107), i.e.

$$\frac{dx}{dt} = \frac{g(t)}{f(x)},$$

is compared with the solution (109), it is seen that the result is obtained more easily, if we incorrectly multiply (107 by the not well-defined factor "$f(x)\,dt$". Then we get

$$(111)\quad f(x)\,dx = g(t)\,dt,$$

which after an integration with respect to x on the left hand side and with respect to t on the right hand side, followed by adding an arbitrary constant on the right hand side, precisely gives (109), i.e.

$$\int f(x)\,dx = \int g(t)\,dt + c.$$

Notice that not every teacher of Calculus will accept equation (111), even if it gives an easy way to remember the solution formula.

5.3 The linear differential equation of first order.

As mentioned in Section 5.1, Introduction, we should be able to solve any linear differential equation of order n by n successive integrations. In principle this is correct, if one is able to guess the missing so-called *"integrating factor"*, followed by the necessary rearrangements of the equation. This is possible for all linear differential equations of first order and for all linear differential equations of constant coefficients of any order n. But when the order is ≥ 2 and the coefficients are variable this is no longer possible in general.

The classical method of solution is the following:

Let

$$(112)\quad \frac{df}{dx} + a(x)f(x) = g(x)$$

be a linear, normalized differential equation of first order. Let

$$A(x) = \int a(x)\,dx$$

be any antiderivative of the coefficient $a(x)$ to $f(x)$ in (112). By a divine inspiration we *guess* the *integrating factor* $\exp(A(x)) \neq 0$. When (112) is multiplied by this factor, we obtain an equivalent equation, which is easily rearranged in the following way,

$$(113)\quad g(x)\,e^{A(x)} = e^{A(x)}\,\frac{df}{dx} + a(x)e^{A(x)}f(x) = e^{A(x)} \cdot \frac{df}{dx} + \frac{d}{dx}\,e^{A(x)} \cdot f(x) = \frac{d}{dx}\left\{e^{A(x)}f(x)\right\},$$

where we have used the rule of differentiation of a product in the opposite direction of the "usual one". Obviously (113) is an ordinary integration problem, so

$$e^{A(x)}f(x) = \int g(x)e^{A(x)}dx + c,$$

and the complete solution is

$$(114)\quad f(x) = e^{-A(x)}\int g(x)e^{A(x)}dx + c \cdot e^{-A(x)} = \varphi(x)\int \frac{g(x)}{\varphi(x)}\,dx + c \cdot \varphi(x),$$

where

$$\varphi(x) = e^{-A(x)} = e^{-\int a(x)\,dx}$$

is a solution of the corresponding homogeneous equation, which is obtained by putting $g(x) \equiv 0$ in (112).

Alternatively we exploit that linear differential equations of first order must be a variant of the usual integration problem. Hence, there must exist a function $\varphi(x) \neq 0$ in the open interval I of consideration, such that the problem can be written in the form

$$(115)\quad \frac{d}{dx}\left\{\frac{f(x)}{\varphi(x)}\right\} = h(x).$$

When (115) is integrated, we get with an arbitrary constant c that

$$\frac{f(x)}{\varphi(x)} = \int h(x)\,dx + c,$$

so the complete solution of (115) is

$$(116)\quad f(x) = \varphi(x) \int h(x)\,dx + c \cdot \varphi(x), \qquad c \in \mathbb{R} \text{ arbitrary}.$$

It follows from the structure of (116) why we have chosen to put $\varphi(x)$ in the denominator of (115) in this analysis. Notice also that $\varphi(x)$ is a solution of the corresponding homogeneous equation.

According to (116) a particular integral of (116) is given by

$$(117)\quad f_0(x) = \varphi(x) \int h(x)\,dx,$$

without the arbitrary constant.

Now, if we instead differentiate the left hand side of (115), we get

$$\frac{1}{\varphi(x)}\,\frac{df}{dx} - \frac{\varphi'(x)}{\varphi(x)^2}\,f(x) = h(x).$$

Since we have assumed that $\varphi(x) \neq 0$ in I, this equation – and hence also (115) – is equivalent to

$$(118)\quad \frac{df}{dx} - \frac{\varphi'(x)}{\varphi(x)}\,f(x) = \varphi(x)h(x).$$

When (118) and (112) are compared, we get the identifications

$$(119)\quad -\frac{\varphi'(x)}{\varphi(x)} = -\frac{d}{dx}\,\ln|\varphi(x)| = a(x) \qquad \text{and} \qquad g(x) = \varphi(x)h(x).$$

It follows from the first equation of (119) that

$$\ln|\varphi(x)| = -\int a(x)\,dx = -A(x), \qquad \text{i.e.} \qquad \varphi(x) = e^{-A(x)},$$

because $\varphi(x) \neq 0$ is continuous, so we can choose $\varphi(x)$ positive. We see that this result agrees with (114).

We furthermore get from (117) that a particular solution is given by

$$(120)\quad f_0(x) = \varphi(x) \int h(x)\, dx = \varphi(x) \int \frac{g(x)}{\varphi(x)}\, dx = e^{-A(x)} \int g(x) e^{A(x)} dx,$$

which is in agreement with (114).

The two methods of solution are therefore equivalent, which they ought to be. The only difference is that the latter method also invites to guess an integration factor. This cannot be said of the former method, where one just puts some functions into a formula and then integrates, no matter how difficult this may be.

In this connection one should also mention a third method. Consider again a differential equation of the form (112),

$$\frac{df}{dx} + a(x)\, f(x) = g(x).$$

If one by a divine inspiration, or by some reasoning on the structure of $a(x)$ and $g(x)$, can guess one particular solution $f_0(x)$, then the task is reduced to finding a nontrivial solution of the corresponding homogeneous equation, i.e.

$$\varphi(x) = e^{-\int a(x)\, dx},$$

because the complete solution is given by

$$f(x) = f_0(x) + c \cdot \varphi(x), \qquad x \in I, \qquad c \in \mathbb{R} \text{ arbitrary.}$$

In elementary courses of Calculus it happens quite often that one can guess such a particular solution without any computation.

Similarly, if one immediately sees that $\varphi(x)$ is a nontrivial solution of the corresponding homogeneous equation, then the task is reduced to the computation of one particular solution given by

$$(121)\quad f_0(x) = \varphi(x) \int \frac{g(x)}{\varphi(x)}\, dx.$$

Example 5.1 We get a very important case when we choose $\varphi(x) = e^{ax}$, where $a \in \mathbb{C}$ is a complex constant. Then equation (120) is written

$$(122)\quad \frac{df}{dx} - a \cdot f(x) = h(x) e^{ax} = g(x),$$

i.e. (122) is a differential equation of first order with constant coefficients. The solution is given by e.g. using (116),

$$(123)\quad f(x) = e^{ax} \int e^{-ax} g(x)\, dx + c \cdot e^{ax}, \qquad c \text{ an arbitrary constant.}$$

Notice that since we already have introduced the complex exponential function

$$e^{(\alpha + \mathrm{i}\,\beta)x} := e^{\alpha x} \cos \beta x + \mathrm{i}\, e^{\alpha x} \sin \beta x,$$

it is easy to check that this example also holds when $a \in \mathbb{C}$ is a complex constant. $\Diamond$

Example 5.2 A similar and almost just as important example is obtained when $\varphi(x) = x^{\alpha}$, $x > 0$. Then equation (118) becomes

$$\frac{df}{dx} - \frac{\alpha}{x}\, f(x) = h(x) \cdot x^{\alpha} = g(x),$$

or, after multiplication with x,

$$(124)\quad x\,\frac{df}{dx} - \alpha\, f(x) = x \cdot g(x).$$

The complete solution of this *Euler differential equation* is given by

$$(125)\quad f(x) = x^{\alpha} \int x^{-\alpha} g(x)\, dx + c \cdot x^{\alpha}, \qquad c \in \mathbb{R} \text{ arbitrary.}$$

If $g(x)$ has the structure of a polynomial, though the exponents need not be integers, then we may expect that some particular solution $f_0(x)$ has a similar structure, so one may *alternatively* spend some time on guessing $f_0(x)$. $\Diamond$

Summary. Given the linear, inhomogeneous differential equation of first order

$$(126)\quad \frac{df}{dx} + a(x)\, f(x) = g(x), \qquad x \in I,$$

where $a(x)$ and $g(x)$ are continuous functions in the open interval I.

1) If we put

$$A(x) = \int a(x)\, dx,$$

(any fixed antiderivative of $a(x)$ will do), then the complete solution of (126) is given by

$$(127)\quad f(x) = e^{-A(x)} \int e^{A(x)} g(x)\, dx + c \cdot e^{-A(x)}, \qquad c \in \mathbb{R} \text{ arbitrary.}$$

2) In some cases it can directly be seen that some $\varphi(x) \neq 0$ in I is a solution of the corresponding homogeneous equation (where $g(x) \equiv 0$). Then the complete solution is given by

$$(128)\quad f(x) = \varphi \int \frac{g(x)}{\varphi(x)}\, dx + c \cdot \varphi(x), \qquad c \in \mathbb{R} \text{ arbitrary.}$$

3) In other cases one may guess by a happy inspiration that $b(x) \neq 0$ in an integrating factor of (126), so

$$b(x)\,\frac{df}{dx} + b(x)a(x)f(x) = \frac{d}{dx}\,\{b(x)f(x)\} = b(x)g(x).$$

In this case the complete solution of (126) is given by

$$(129)\quad f(x) = \frac{1}{b(x)} \int b(x)g(x)\, dx + \frac{c}{b(x)}, \qquad c \in \mathbb{R} \text{ arbitrary.}$$

When (128) and (129) are compared we see that $\varphi(x) = \dfrac{1}{b(x)}$.

4) Finally, it is sometimes possible to *guess* a particular solution $f_0(x)$. In this case one shall perform one integration operation less than in the official solution formula, which is saving time. The complete solution is given by

(130) $$f(x) = f_0(x) + c \cdot e^{-\int a(x)\, dx}, \qquad c \in \mathbb{R} \text{ arbitrary.}$$

Example 5.3 Consider the differential equation

$$\frac{df}{dx} - \tan x \cdot f(x) = 2 \sin x, \qquad x \in \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[,$$

which we solve in three different ways.

1) Since $a(x) = -\tan x$, we have

$$A(x) = -\int \tan x\, dx = -\int \frac{\sin x}{\cos x}\, dx = \ln \cos x,$$

where we have used that $\cos x > 0$ in the given interval.

Then, according to the solution formula (127),

$$\begin{aligned}f(x) &= e^{-A(x)}\int e^{A(x)}g(x)\,dx + c\cdot e^{-A(x)}\\ &= \frac{1}{\cos x}\int 2\sin x\cdot\cos x\,dx + \frac{c}{\cos x}\\ &= -\cos x + \frac{c}{\cos x}, \qquad c\in\mathbb{R}\text{ arbitrary.}\end{aligned}$$

2) *Alternatively*, it is not hard to see that $\cos x > 0$ is an integrating factor, so

$$2\sin x\cdot\cos x = -\frac{d}{dx}\left\{\cos^2 x\right\} = \cos x\cdot\frac{df}{dx} - \sin x\cdot f(x) = \frac{d}{dx}\{\cos x\cdot f(x)\}.$$

Thus by integration,

$$\cos x\cdot f(x) = c - \cos^2 x,$$

and the complete solution is given by

$$f(x) = -\cos x + \frac{c}{\cos x}, \qquad c\in\mathbb{R}\text{ arbitrary.}$$

3) If $f_0(x) = \cos x$ is inserted into the left hand side of the differential equation, we get

$$\frac{df_0}{dx} - \tan x\cdot f(x) = -\sin x - \tan x\cdot\cos x = -2\sin x.$$

Hence, $f_0(x) = -\cos x$ is a particular solution.

In our next guess we choose $f(x) = \dfrac{1}{\cos x}$. By insertion into the left hand side of the differential equation we get

$$\frac{df}{dx} - \tan x\cdot f(x) = \frac{\sin x}{\cos^2 x} - \frac{\sin x}{\cos x}\cdot\frac{1}{\cos x} = 0,$$

so $\varphi(x) = 1/\cos x$, $x\in\,]-\pi/2,\pi/2[$, is a solution of the corresponding homogeneous equation. The complete solution is

$$f(x) = -\cos x + \frac{c}{\cos x}, \qquad c\in\mathbb{R}\text{ arbitrary.} \qquad \Diamond$$

5.4 Linear differential equations of constant coefficients.

Case n = 1. This was already handled in Example 5.1 of Section 5.3.
If $a\in\mathbb{C}$ is a constant, and

$$(131)\quad \frac{df}{dx} + a\,f(x) = g(x), \qquad x\in I,$$

then the complete solution is given by

$$f(x) = e^{-ax}\int e^{ax}g(x)\,dx + c\cdot e^{-ax}, \qquad x\in I,$$

where $c \in \mathbb{R}$, or $c \in \mathbb{C}$, is an arbitrary constant.

Case n =2. We give here a totally different treatment of second order differential equations of constant coefficients compared with the usual one presented in the elementary courses of Calculus.

Let a and b be (complex) constants. Then consider for any given continuous function $g(x)$ in the open interval I the special differential equation of second order,

(132) $$\frac{d}{dx}\left\{e^{(a-b)x}\,\frac{d}{dx}\left(e^{-ax}f(x)\right)\right\} = g(x), \qquad x \in I.$$

Obviously, (132) has been constructed so that it can easily be solved by two integrations with suitable rearrangements in between. We shall later do this in all details, but first we shall compute the left hand side of (132) in order to specify the type of differential equations which are solved by means of (132).

When the left hand side of (132) is computed, we obtain

$$\begin{aligned}
&\frac{d}{dx}\left\{e^{(a-b)x}\,\frac{d}{dx}\left(e^{-ax}f(x)\right)\right\} = \frac{d}{dx}\left\{e^{(a-b)x}\left(e^{-ax}\,\frac{df}{dx} - a\,e^{-ax}f(x)\right)\right\}\\
&\quad = \frac{d}{dx}\left\{e^{-bx}\,\frac{df}{dx} - a\,e^{-bx}f(x)\right\} = e^{-bx}\,\frac{d^2f}{dx^2} - (a+b)e^{-bx}\,\frac{df}{dx} + ab\,e^{-bx}f(x)\\
&\quad = e^{-bx}\left\{\frac{d^2f}{dx^2} - (a+b)\,\frac{df}{dx} + ab\cdot f(x)\right\} = g(x).
\end{aligned}$$

Since $e^{-bx} \neq 0$, it follows that (132) is equivalent to

(133) $$\frac{d^2f}{dx^2} - (a+b)\,\frac{df}{dx} + ab\cdot f(x) = e^{bx}g(x) = h(x), \qquad x \in I,$$

hence

(134) $$g(x) = e^{-bx}h(x), \qquad x \in I,$$

for given $h(x)$.

It follows that (133) is a linear, inhomogeneous differential equation of second order with constant coefficients and that any such differential equation can be written in the form (133) for some constants a, b and some function $h(x)$.

The corresponding *characteristic polynomial* is obtained by the left hand side of (133) by the correspondences

$$\frac{d^2f}{dx^2} \longleftrightarrow R^2, \qquad \frac{df}{dx} \longleftrightarrow R, \qquad f \longleftrightarrow 1,$$

i.e.

(135) $$R^2 - (a+b)R + ab = (R-a)(R-b),$$

and we see that the differential equation (132) was constructed in such a way that the characteristic polynomial has the two roots a and b.

Notice that since (132) is solved by two integrations, we must have two linearly independent sets of solutions to the corresponding homogeneous equation,

(136) $$\frac{d}{dx}\left\{e^{(a-b)x}\,\frac{d}{dx}\left(e^{-ax}f(x)\right)\right\}=0.$$

Obviously, $f_1(x)=e^{ax}$ is a solution, because already

$$\frac{d}{dx}\left\{e^{-ax}f_1(x)\right\}=\frac{d1}{dx}=0.$$

Similarly, when $b\neq a$, we get that $f_2(x)=e^{bx}$ is a linearly independent solution of the homogeneous equation.

When $b=a$, equation (136) is reduced to

$$\frac{d^2}{dx^2}\left\{e^{-ax}f(x)\right\}=0,$$

which has the complete solution

$$f(x)=c_1e^{ax}+c_2x\,e^{ax},\qquad c_1,\,c_2\in\mathbb{R}\quad(\text{or }\mathbb{C})\quad\text{arbitrary constants.}$$

Since (132) and (133) are equivalent, these equations have the same complete solution. Hence we have proved

Theorem 5.1 *Consider the linear homogeneous differential equation of second order*

$$\frac{d^2f}{dx^2}-(a+b)\,\frac{df}{dx}+ab\,f(x)=0,$$

where the characteristic polynomial has the two (complex) roots a and b.

1) If $a\neq b$, then the complete solution is

$$f(x)=c_1e^{ax}+c_2e^{bx},\qquad c_1,\,c_2\ \textit{arbitrary constants.}$$

2) If $a=b$, then the complete solution is

$$f(x)=c_1e^{ax}+c_2x\,e^{ax},\qquad c_1,\,c_2\ \textit{arbitrary constants.}$$

Example 5.4 Since we already have introduced the complex exponential function,

$$e^{\mathrm{i}\,x}:=\cos x+\,\mathrm{i}\,\sin x,\qquad \cos x=\frac{1}{2}\left\{e^{\mathrm{i}\,x}+e^{-\mathrm{i}\,x}\right\},\qquad \sin x=\frac{1}{2\,\mathrm{i}}\left\{e^{\mathrm{i}\,x}-e^{-\mathrm{i}\,x}\right\},$$

we can allow characteristic polynomials with complex roots.

An important example is

$$\frac{d^2f}{dx^2}+f(x)=0.$$

The characteristic polynomial R^2+1 has the two simple roots $R=\pm\,\mathrm{i}$, so according to Theorem 5.1 the complete solution is

$$f(x)=\tilde{c}_1e^{\mathrm{i}\,x}+\tilde{c}_2e^{-\mathrm{i}\,x}=c_1\cos x+c_2\sin x,\qquad c_1,\,c_2\ \text{arbitrary.}$$

It is easy to check that these are all solutions. What is new here is that we have *proved* that they are all the possible solutions. $\Diamond$

Let us return to the solution of the inhomogeneous equation (132), i.e.

$$(137)\quad \frac{d}{dx}\left\{e^{(a-b)x}\,\frac{d}{dx}\left(e^{-ax}f(x)\right)\right\}=g(x).$$

Since this equation is linear, and since all the solutions of the homogeneous equation are given by Theorem 5.1, we only have to find one particular solution of the inhomogeneous equation, because then the complete solution is obtained by adding all the solutions of Theorem 5.1 to the particular solution. Hence, it is no need to add the arbitrary constants when we integrate (137).

When this is done we get by a rearrangement that

$$\frac{d}{dx}\left\{e^{-ax}f(x)\right\}=e^{-(b-a)x}\int g(x)\,dx.$$

By another integration and another rearrangement we obtain the particular solution

$$(138)\quad f(x)=e^{ax}\int e^{(b-a)x}\left\{\int g(x)\,dx\right\}dx=e^{ax}\int e^{(b-a)x}\left\{\int e^{-bx}h(x)\,dx\right\}dx.$$

If one does not like the two successive integrations in (138), one can *alternatively* perform a partial integration. Here we have two cases:

1) When $b \neq a$, then

$$\begin{aligned} f(x) &= e^{ax}\left\{\frac{1}{b-a}\,e^{(b-a)x}\int g(x)\,dx - \frac{1}{b-a}\int e^{(b-a)x}g(x)\,dx\right\} \\ &= \frac{1}{b-a}\,e^{bx}\int g(x)\,dx - \frac{1}{b-a}\,e^{ax}\int e^{(b-a)x}g(x)\,dx \\ (139) \qquad &= \frac{1}{b-a}\,e^{bx}\int e^{-bx}h(x)\,dx - \frac{1}{b-a}\,e^{ax}\int e^{-ax}h(x)\,dx. \end{aligned}$$

2) When $b = a$, we get instead

$$\begin{aligned} f(x) &= e^{ax}\int 1\cdot\left\{\int g(x)\,dx\right\}dx = e^{ax}\left\{x\int g(x)\,dx - \int s\,g(x)\,dx\right\} \\ (140) \qquad &= x\,e^{ax}\int e^{-ax}h(x)\,dx - e^{ax}\int x\,e^{-ax}h(x)\,dx. \end{aligned}$$

We have proved

Theorem 5.2 *Consider the linear inhomogeneous differential equation of second order with constant coefficients*

$$(141) \quad \frac{d^2 f}{dx^2} - (a+b)\,\frac{df}{dx} + ab\,f(x) = h(x), \qquad x \in I,$$

where the characteristic polynomial $R^2 - (a+b)R + ab$ has the two complex roots a and b.

A particular solution is given by

$$(142) \quad f_0(x) = e^{ax}\int e^{(b-a)x}\left\{\int e^{-bx}h(x)\,dx\right\}dx.$$

Alternatively*, (142) is written*

1) For $b \neq a$,

$$(143) \quad f_0(x) = \frac{1}{b-a}\,e^{bx}\int e^{-bx}h(x)\,dx - \frac{1}{b-a}\,e^{ax}\int e^{-ax}h(x)\,dx.$$

2) For $b = a$,

$$(144) \quad f_0(x) = x\,e^{ax}\int e^{-ax}h(x)\,dx - e^{ax}\int x\,e^{-ax}h(x)\,dx.$$

The complete solution of (141) is either given by (142), where we add an arbitrary constant after each integration, or alternatively

1) For $b \neq a$,

$$(145) \quad f(x) = \frac{1}{b-a}\left\{e^{bx}\int e^{-bx}h(x)\,dx - e^{ax}\int e^{-ax}h(x)\,dx\right\} + c_1 e^{ax} + c_2 e^{bx},$$

where c_1 and c_2 are arbitrary constants.

2) For $b = a$,

$$(146)\quad f(x) = x\,e^{ax}\int e^{-ax}h(x)\,dx - e^{ax}\int x\,e^{-ax}h(x)\,dx + c_1e^{ax} + c_2x\,e^{ax},$$

Remark 5.1 It is often possible to *guess* a particular solution $f_0(x)$ by analyzing the right hand side $h(x)$. If this procedure is successful, the solution formula is reduced to

1) For $b \neq a$,

$$f(x) = f_0(x) + c_1e^{ax} + c_2e^{bx}.$$

2) For $b = a$,

$$f_0(x) + c_1e^{ax} + c_2x\,e^{ax}. \qquad \Diamond$$

Example 5.5 For linear differential equations the structure of the set of solutions is given by adding all solutions of the homogeneous equation to one particular solution. Therefore, by guessing one particular solution one can often reduce the computations considerably.

Let us consider the inhomogeneous differential equation

$$(147)\quad \frac{d^2f}{dx^2} + f(x) = x.$$

The characteristic polynomial $R^2 + 1$ has the two simple complex roots, $a = \mathrm{i}$ and $b = -\,\mathrm{i}$, and we have previously seen in Example 5.4 that the complete solution of the homogeneous equation is given by

$$\tilde{c}_1e^{\mathrm{i}\,x} + \tilde{c}_2e^{-\mathrm{i}\,x} = c_1\cos x + c_2\sin x.$$

If we use the solution formula (142), we get the following expression of a particular solution,

$$f_0(x) = e^{\mathrm{i}\,x}\int e^{-2\,\mathrm{i}\,x}\left\{\int e^{\mathrm{i}\,x}x\,dx\right\}dx,$$

which according to (143) also can be written

$$f_0(x) = -\frac{1}{2\,\mathrm{i}}\,e^{-\mathrm{i}\,x}\int e^{\mathrm{i}\,x}x\,dx + \frac{1}{2\,\mathrm{i}}\,e^{\mathrm{i}\,x}\int e^{-\mathrm{i}\,x}x\,dx.$$

It is possible (see below) to compute both expressions, but it is of course much easier to realize that $f_0(x) = x$ is trivially a particular solution of (147), so the complete solution is

$$f(x) = x + c_1\cos x + c_2\sin x.$$

For comparison we see that since

$$\int x\,e^{ax}\,dx = \frac{1}{a}\,x\,e^{ax} - \frac{1}{a^2}\,e^{ax}\text{for } a \neq 0,$$

it follows from the latter of the two solution formulæ that

$$\begin{aligned} f_0(x) &= -\frac{1}{2\,\mathrm{i}}\,e^{-\mathrm{i}\,x}\int_{a=\mathrm{i}} x\,e^{\mathrm{i}\,x}dx+\frac{1}{2\,\mathrm{i}}\,e^{\mathrm{i}\,x}\int_{a=-\,\mathrm{i}} x\,e^{-\,\mathrm{i}x}dx \\ &= -\frac{1}{2\,\mathrm{i}}\,e^{-\mathrm{i}\,x}\left\{\frac{1}{\mathrm{i}}\,x\,e^{\mathrm{i}\,x}-\frac{1}{\mathrm{i}^2}\,e^{\mathrm{i}\,x}\right\}+\frac{1}{2\,\mathrm{i}}\,e^{\mathrm{i}\,x}\left\{-\frac{1}{\mathrm{i}}\,x\,e^{-\mathrm{i}\,x}-\frac{1}{\mathrm{i}^2}\,e^{-\mathrm{i}\,x}\right\} \\ &= -\frac{1}{2\,\mathrm{i}}\left\{\frac{1}{\mathrm{i}}\,x+1\right\}+\frac{1}{2\,\mathrm{i}}\left\{-\frac{1}{\mathrm{i}}\,x+1\right\}=\left\{\frac{1}{2}\,x-\frac{1}{2\,\mathrm{i}}\right\}+\left\{\frac{1}{2}\,x+\frac{1}{2\,\mathrm{i}}\right\}=x, \end{aligned}$$

so we get the same solution after some fairly long computations. $\Diamond$

Example 5.6 It is often possible to guess a particular solution even when the theory does not explicitly give a suggestion of such a solution. Consider e.g. the equation

$$(148)\quad \frac{d^2f}{dx^2}-2\,\frac{df}{dx}+f(x)=e^x.$$

The characteristic polynomial $R^2-2R+1=(R-1)^2$ has $R=1$ as a root of multiplicity 2, so the corresponding homogeneous equation has the complete solution

$$c_1e^x+c_2x\,e^x.$$

Since the right hand side e^x of (148) already is a solution of the homogeneous equation, it is no need to guess e^x, because we will automatically get $0\neq e^x$. The same can be said of $x\cdot e^x$. The trick is now to guess a particular solution of the form

$$(149)\quad f_0(x)=c\cdot x\cdot x\,e^x=c\,x^2e^x,$$

which is not a solution of the homogeneous equation.

Since

$$\frac{df_0}{dx}=c\,x^2e^x+2c\,x\,e^x,\qquad \frac{d^2f_0}{dx^2}=c\,x^2e^x+4c\,x\,e^x+2c\,e^x,$$

we get by insertion in the left hand side of (148) that

$$\frac{d^2f_0}{dx^2}-2\,\frac{df_0}{dx}+f_0(x)=c\,\{-2+1\}x^2e^x+c\{4-4\}x\,e^x+2c\,e^x=2c\,e^x.$$

so we get a particular solution by choosing $c=\dfrac{1}{2}$.

If instead the right hand side is $x\,e^x$, the idea is to guess

$$f_1(x)=c_1x^3e^x+c_2x^2e^x,$$

and then find the right constants c_1 and c_2. The details are left to the reader. $\Diamond$

The case n > 2. When a linear differential equation with constant coefficient has order $n>2$, the method of solution is the same as above. It is left to the reader to prove the following theorem:

Theorem 5.3 *Consider the linear, inhomogeneous, normalized differential equation of order $n > 2$ with constant coefficients*

$$(150)\quad \frac{d^n f}{dx^n} + a_{n-1}\,\frac{d^{n-1} f}{dx^{n-1}} dx^{n-1} + \cdots + a_1\,\frac{df}{dx} + a_0 f(x) = h(x).$$

Let $\alpha_1, \dots, \alpha_n$ be the n complex roots of the characteristic polynomial, i.e.

$$R^n + a_{n-1}R^{n-1} + \cdots + a_1 R + a_0 = (R-\alpha_1)(R-\alpha_2)\cdots(R-\alpha_n).$$

Then a particular solution is given by the integration formula

$$(151)\quad f_0(x) = e^{\alpha_1 x}\int e^{(\alpha_2-\alpha_1)x}\left\{\int e^{(\alpha_3-\alpha_2)x}\left(\cdots\left\{\int e^{-\alpha_n x} h(x)\,dx\right\}\cdots dx\right)dx\right\}dx,$$

where any ordering of the roots $\alpha_1, \dots, \alpha_n$ can be used.

If one adds an arbitrary constant c_j after the j-th integration, then (151) gives the complete solution of (150).

Remark 5.2 An easy way of setting up (151) is to enumerate the roots of the characteristic polynomial in increasing order,

$$\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n .$$

Note that

$$\alpha_1 + (\alpha_2 - \alpha_1) + (\alpha_3 - \alpha_2) + \cdots + (\alpha_n - \alpha_{n-1}) - \alpha_n = 0,$$

which may be used of a check of (151) in specific cases. $\Diamond$

Remark 5.3 The complete solution of the corresponding homogeneous equation can *alternatively* be found in the following way:

If α is a root of the characteristic polynomial of multiplicity k, then this particular root contributes to the solution of the homogeneous equation by

$$c_{\alpha,0} e^{\alpha x} + c_{\alpha,1} x \, e^{alphax} + \cdots + c_{\alpha,k-1} x^{k-1} e^{\alpha x}.$$

When this is done for every of the different roots we obtain the complete solution of the corresponding homogeneous solution by adding all these, where the $c_{\alpha,j}$ of course are arbitrary constants. $\Diamond$

In general, (151) is very complicated to compute, so a good advice is always first to try to *guess* a particular solution by analyzing the right hand side of the differential equation.

5.5 Euler's differential equation.

This equation is no more compulsory in the first year of Calculus, so we shall be very brief here.

By *Euler's differential equation* of order n we understand a differential equation of the form

$$x^n \frac{d^n f}{dx^n} + a_{n-1} x^{n-1} \frac{d^{n-1} f}{dx^{n-1}} + \cdots + a_j x^j \frac{d^j f}{dx^j} + \cdots + a_1 x \frac{df}{dx} + a_0 f(x) = g(x), \qquad x > 0,$$

where $a_0, \dots, a_{n-1}$ are constants, and $g(x)$ is a given function.

The usual method of solution is to use the change of variables $x = e^t$. Then *Euler's differential equation* can be transformed into an inhomogeneous differential equation of order n and constant coefficients (NB, *not the same constants as in the Euler differential equation itself*),

$$\frac{d^n F}{dt} + b_{n-1} \frac{d^{n-1} F}{dt^{n-1}} + \cdots + b_1 \frac{dF}{dt} + b_0 F(t) = G(t), \qquad t \in \mathbb{R},$$

which can be solved by the methods in Section 5.4. When the complete solution of this equation if found, we get the complete solution of Euler's differential equation by the correspondence

$$f(x) := F(\ln x),$$

using that $t = \ln x$. In particular, if a is a real root of multiplicity 1 of the characteristic equation of the corresponding t-equation, then

$$e^{at} = x^a$$

is a solution of the homogeneous Euler equation, and if a is of multiplicity 2, then we also have that

$$t\,e^{at} = \ln x \cdot x^a,$$

etc.

If the Euler equation has real coefficients and $a = \alpha + \mathrm{i}\,\beta$ and $\overline{a} = \alpha - \mathrm{i}\,\beta$ are complex conjugated roots of the characteristic polynomial, we know already that the corresponding solution of homogeneous equation can be written

$$c_1 e^{\alpha t}\cos(\beta t) + c_2 e^{\alpha t}\sin(\beta t) = c_1 x^{\alpha}\cos(\beta\,\ln x) + c_2 x^{\alpha}\sin(\beta\,\ln x),$$

etc..

There is therefore a shortcut concerning the solutions of the homogeneous Euler differential equation: Insert formally $f(x) = x^R$ into the left hand side of Euler's differential equation and put it equal to 0. Since

$$x^j \frac{d^j}{dx^j} x^R = c_{j,R} \cdot x^R,$$

where $c_{j,R}$ is some constant, depending on R and the exponent R on the right hand side does not depend on j, it can be shown that we in this way get precisely the characteristic equation of the corresponding t-equation. By using the interpretations above of x^a etc., where a is one of the roots, it is quite easy to find the solution of the homogeneous Euler equation.

5.6 Linear differential equations of second order with variable coefficients.

This section has only been added for completeness. One may expect that it is first taught in Calculus on the second year. It may, however, be of some help for those students who want to understand how some second order equations with variable coefficients are solved in Physics.

We consider a normalized linear inhomogeneous differential equation of second order with variable coefficients

$$(152) \quad \frac{d^2 f}{dx^2} + a(x)\, \frac{df}{dx} + b(x)\, f(x) = h(x), \qquad x \in I,$$

where $a(x)$, $b(x)$ and $h(x)$ are given continuous functions defined in an open interval I.

An equation is called *normalized*, when the coefficient of the highest order term, here $\dfrac{d^2 f}{dx^2}$, is 1 like in (152). Notice that the most common solution formulæ, whenever they exist (!), refer to the *normalized* equation. It is therefore a good strategy with the exception of *Euler's differential equation* considered in Section 5.5, always to start by normalizing the differential equation, i.e. divide it by the coefficients of the highest order term.

Unlike the equations in the previous sections, equation (152) does not in general have a solution formula, and this is one of the reasons why it usually is not included in the courses of the first year in Calculus.

It is worth to analyze the problem of equation (152). As mentioned earlier several times it should in principle be possible to produce a solution formula including two successive integrations, because the equation is linear of second order. An equivalent statement is that it should be possible to write (152) in the form of a system of two linear differential equations of first order, by introducing a new unknown function u,

$$(153) \quad \begin{cases} \dfrac{df}{dx} - q(x)\, f(x) = u(x), \\[2ex] \dfrac{du}{dx} - p(x)\, u(x) = h(x), \end{cases}$$

where the task is to find the unknown coefficients $q(x)$ and $p(x)$, such that the complete solution of $f(x)$ is the same in both (152) and (153).

By eliminating u of (153) we obtain after quite some computation that

$$(154) \quad \frac{d^2 f}{dx^2} - \{p(x) + q(x)\}\, \frac{df}{dx} + \left\{ p(x) \cdot q(x) - \frac{dq}{dx} \right\} f(x) = h(x).$$

By assumption, (154) and (152) are the same equation, hence by identification of the coefficients,

$$(155)\quad p(x)+q(x)=-a(x) \qquad \text{and} \qquad p(x)\cdot q(x)-\frac{dq}{dx}=b(x).$$

By eliminating $p(x)$ of (155) we get a so-called *Riccati equation* in $q(x)$,

$$(156)\quad \frac{dq}{dx}+a(x)\,q(x)+q(x)^2=-b(x), \qquad x\in I.$$

It can be shown by using the existence theorem for ordinary differential equations that (156) does have solutions $q(x)$. According to the first equation of (155) we then get a unique $p(x)=-a(x)-q(x)$, and we have proved that (152) *in principle* can be reformulated as (153). Hence, the solution of (152) can in fact be found by two integrations with suitable rearrangements of the equation in between.

We also get two arbitrary constants as coefficients of two linearly independent solutions of the corresponding homogeneous equation.

When $p(x)$ and $q(x)$ are given, it follows from (153) that

$$u(x)=e^{\int p(x)\,dx}\int e^{-\int p(x)\,dx}h(x)\,dx+c_2e^{\int p(x)\,dx},$$

which inserted into

$$\frac{df}{dx}-q(x)\,f(x)=u(x),$$

followed by the usual solution procedure from Section 5.3, gives a rather messy solution formula. We shall not give it here because this might mislead the reader to believe that this is the standard procedure, which it is *not!*

Let us return to the *Riccati differential equation* (156). It has been known for more than a century that there *exists no general solution formula for (156)!* But if just one solution $q_0(x)$ of (156) is known, then it can be totally solved. The only way of finding an exact solution at this stage is to make a qualified guess by some clever inspiration, which is hardly a solution procedure one would demand of students in their first year of Calculus.

This unfortunate fact propagates back to equation (152), for if it existed we would also have a solution formula for the Riccati equation (156), which we do not have. But .... *If one knows just one solution $\varphi(x)\neq 0$ for $x\in I$ of the corresponding homogeneous equation, then one can find the complete solution of equation (152).*

We shall now prove the latter claim.

Suppose that $\varphi(x)\neq 0$ for $x\in I$ is a *given* solution of the corresponding homogeneous equation. Then there must be a way of writing (152) in the equivalent form

$$(157)\quad \frac{d}{dx}\left\{\frac{1}{\psi(x)}\,\frac{d}{dx}\left(\frac{1}{\varphi(x)}\,f(x)\right)\right\}=g(x), \qquad x\in I,$$

because it is trivial that $f(x)=\varphi(x)$ is a solution of the corresponding homogeneous equation. We only need to find $\psi(x)\neq 0$ and $g(x)$ in I, so that (152) and (157) have the same set of solutions.

Assume for the time being that both $\psi(x)$ and $g(x)$ have been found. Then we get by an integration of (157) followed by a rearrangement that

$$\frac{d}{dx}\left\{\frac{1}{\varphi(x)}\,f(x)\right\} = \psi(x)\int g(x)\,dx + c_2\psi(x),$$

hence by repeating this procedure,

$$(158)\quad)f(x) = \varphi(x)\int \psi(x)\left\{\int g(x)\,dx\right\}dx + c_1\varphi(x) + c_2\varphi(x)\int \psi(x)\,dx.$$

We see that the special form (157) of the differential equation gives the simple structure (158) of the solutions.

If we compute the left hand side of (157) and multiply the result by $\varphi(x)\psi(x)$, we get after some reduction that (157) is equivalent to

$$(159)\quad \frac{d^2 f}{dx^2} - \left\{\frac{\pi'(x)}{\psi(x)} + 2\,\frac{\varphi'(x)}{\varphi(x)}\right\}\frac{df}{dx} + \frac{\varphi'(x)}{\varphi(x)}\left\{\frac{\psi'(x)}{\psi(x)} + 2\,\frac{\varphi'(x)}{\varphi(x)} - \frac{\varphi''(x)}{\varphi'(x)}\right\}f(x) = \varphi(x)\psi(x)f(x).$$

Since (159) by assumption has the same set of solutions as (152), these two equations must be identical. By identifying the coefficients and terms we get in particular

$$(160)\quad \begin{cases} \dfrac{\psi'(x)}{\psi(x)} + 2\,\dfrac{\varphi'(x)}{\varphi(x)} = \dfrac{d}{dx}\,\ln\left\{\psi(x)\varphi(x)^2\right\} = -a(x), \\[2ex] \varphi(x)\psi(x)g(x) = h(x). \end{cases}$$

It follows from the first equation of (160) that

$$(161)\quad \psi(x)\varphi(x)^2 = e^{-\int a(x)\,dx} := W_0(x).$$

The notation $W_0(x)$ shall indicate that (161) is one of the so-called *Wrońskians* corresponding to (152), i.e. traditionally,

$$W(x) = \varphi_1\,\frac{d\varphi_2}{dx} - \varphi_2\,\frac{d\varphi_1}{dx} = \begin{vmatrix} \varphi_1(x) & \varphi_2(x) \\ \varphi_1'(x) & \varphi_2'(x) \end{vmatrix} = e\cdot e^{-\int a(x)\,dx}, \qquad c \neq 0 \text{ arbitrary},$$

where $\varphi_1(x)$ and $\varphi_2(x)$ form a pair of linearly independent solutions of the corresponding homogeneous equation. Obviously (161) shoes that the Wrońskian can always be found.

By using (158) we see that $\varphi_1(x) := \varphi(x)$ together with

$$(162)\quad \varphi_2(x) := \varphi(x)\int \psi(x)\,dx = \varphi(x)\int \frac{1}{\varphi(x)^2}\,e^{-\int a(x)\,dx}dx = \varphi(x)\int\int \frac{W_0(x)}{\varphi(x)^2}\,dx$$

form a pair of linearly independent solutions of the homogeneous differential equation, and a small computation shows that the Wrońskian of this particular pair (φ_1, φ_2) is precisely $W_0(x)$.

Finally, a particular solution is according to (158) given by

$$\begin{aligned} f_0(x) &= \varphi(x)\int \psi(x)\left\{\int g(x)\,dx\right\}dx \\ &= \varphi\int \frac{W_0(x)}{\varphi(x)^2}\left\{\int \frac{\varphi(x)}{W_0(x)}\cdot h(x)\,dx\right\}dx. \end{aligned}$$

Thus we have proved

Theorem 5.4 *We consider the normalized linear inhomogeneous differential equation*

$$(163)\quad \frac{d^2 f}{dx^2} + a(x)\,\frac{df}{dx} + b(x)\,f(x) = h(x), \qquad x \in I,$$

of second order, where $a(x)$, $b(x)$ and $h(x)$ are continuous in the open interval I.

There is in general no solution formula of (163), but if we know just non nontrivial solution $\varphi(x) \neq 0$ for $x \in I$ of the corresponding homogeneous equation, then we can find the complete solution of (163) in the following way:

1) Compute the Wrońskian

$$W_0(x) = \exp\left(-\int a(x)\,dx\right).$$

2) Any solution of the corresponding homogeneous equation can be written as a superposition $a_1\varphi_1(x) + a_2\varphi_2(x)$ of the two functions

$$\varphi_1(x) = \varphi(x) \qquad \text{and} \qquad \varphi_2(x) = \varphi(x)\int \frac{W_0(x)}{\varphi(x)^2}\,dx.$$

3) One particular solution of (163) is given by the formula

$$f_0(x) = \varphi(x)\int \frac{W_0(x)}{\varphi(x)^2}\left\{\int \frac{\varphi(x)\,h(x)}{W_0(x)}\,dx\right\}dx.$$

4) The complete solution of (163) is given by

$$f(x) = f_0(x) + c_1\varphi_1(x) + c_2\varphi_2(x), \qquad x \in I,$$

where c_1 and c_2 are arbitrary constants.

Remark 5.4 Since the two successive integrations in the formula of a particular solution of (163) may be rather complicated to perform, a good advice is always first to analyze (163) in order to see if is is possible to *guess* $f_0(x)$ directly without any integration. $\Diamond$

As mentions earlier the material of this section does no longer belong to a first course in Calculus, but it may be useful for the applications in e.g. Physics, where one needs to solve linear differential equations with or without constant coefficients.

6 Approximations of Functions

6.1 Introduction.

So far we have only used functions like polynomials, logarithms, exponentials, trigonometric and hyperbolic functions and the inverse trigonometric and hyperbolic functions. Even though these functions can describe quite a lot of phenomena they are far from adequate in practical applications. The above mentioned classes of functions are only the simplest ones which every mathematician and engineer should have some knowledge of. For the same reason the teaching of Calculus is mostly focussed only on these functions, thus risking that the student may believe that all other functions may not be so important, which is wrong.

It is of course possible to enlarge the set of functions which may be considered as "known". If one e.g. in Physics continues to encounter the same type of differential equations, then it would be quite reasonable once and for all to find the standard form of this equation and thereby the corresponding standard solutions.

The student probably first met this phenomenon in high school where the complete solution of

$$\frac{d^2 f}{dx^2} + a^2 f(x) = 0, \qquad x \in \mathbb{R}, \qquad a \in \mathbb{R}_+ \text{ a constant,}$$

can be written as a linear combination of the functions $\cos ax$ and $\sin ax$. Since $a \in \mathbb{R}_+$ is only a scaling constant, the generic equation is

$$\frac{d^2 f}{dx^2} + f(x) = 0, \qquad x \in \mathbb{R},$$

with the two linearly independent generic solutions $\cos x$ and $\sin x$.

The example above is so simple that one hardly sees the problem. But if we take a *Bessel differential equation* (no need here to give its precise definition), which in particular is closely connected with problems in polar coordinates in the plane, then we have to define and study some new functions, called *Bessel functions*. In this case the set of "known" functions should for convenience be enlarged by these new Bessel functions, which may be of use for some of the readers, though not for all of them.

This procedure may work to some extent, but unfortunately we shall never be able to exhaust the set of all useful functions in this way.

One problem is that we by this procedure in most cases only will enlarge our set of functions with solutions of *linear* differential equations, because it is very difficult, if possible, to find any generic characteristic of a class of useful nonlinear differential equations. It was, however, realized for over a century ago that in principle any physical process in *nonlinear*, so our arsenal of "known" functions can never be adequate, concerning what we would call *exact solutions*.

A typical and simple example is the equation of a *mathematical pendulum* (which already is a simplified model of the physical situation), where we consider a material point of pass m (chosen conveniently as $m = 1$, since it is cancelled in the derivation of the differential equation), suspended on a string of length ℓ.

We assume that the movement takes place in a plane and that φ is the angle between the string and the vertical axis. It is not difficult to derive that the movement is described by the following *nonlinear*

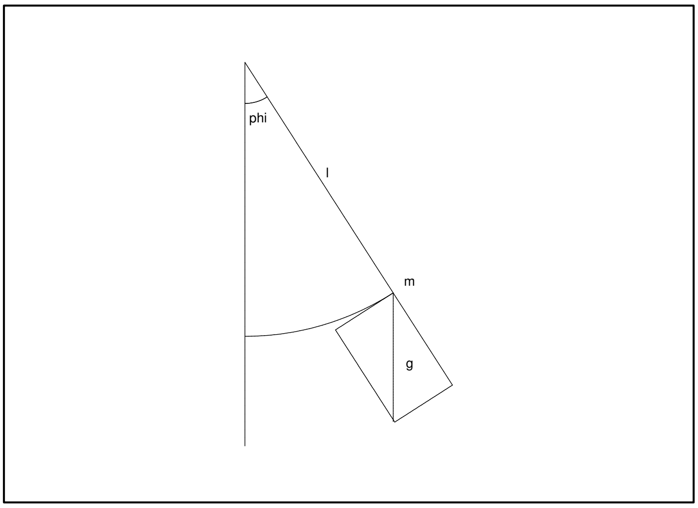

Figure 25: The mathematical pendulum.

differential equation of second order in φ,

(164) $$\frac{d^2\varphi}{dt^2} = -\frac{g}{\ell}\,\sin\varphi.$$

This equation does not have a solution which can be described by our known functions. Fortunately, we are mostly only interested in an approximate solution of an error smaller than a given $\varepsilon > 0$. These considerations from Physics are some of the starting points of the approximation theory, of which we shall only go through the simplest cases.

As a spin-off we obtain some new methods of finding limits and asymptotes of curves, and approximate values of integrals.

Finally, we shall at the end of this chapter return to an approximation of a solution of the mathematical pendulum (164).

6.2 ε - functions.

By $\varepsilon(x)$ we shall understand any function $f(x)$ for which $f(x) \to 0$ for $x \to 0$.

More precisely, to every $\varepsilon > 0$ there exists a $\delta > 0$, such that

(165) $|f(x)| < \varepsilon$, whenever $|x| < \delta$ in the domain of f,

where we furthermore assume that such an x exists for every $\delta > 0$.

The domain of f needs not be a neighbourhood of 0, though it usually is. We only require that 0 is a contact point of the domain of f. One example is when 0 is the endpoint of the domain of the function f.

Frankly speaking, $\varepsilon(x)$ is the set of all such functions, but it is practice to let it denote any function with this property. Since $\varepsilon(x)$ may denote different functions in the same equation, we notice the following rules,

$$(166)\quad \begin{cases} \varepsilon(x) \pm \varepsilon(x) = \varepsilon(x), & \\ a \cdot \varepsilon(x) = \varepsilon(x) & \text{for } a \in \mathbb{R} \text{ constant,} \\ \varepsilon\left(x^n\right) = \varepsilon(x), & \text{for } n \in \mathbb{N} \text{ constant,} \\ g(x) \cdot \varepsilon(x) = \varepsilon(x), & \text{for } g(x) \text{ bounded in the neighbourhood of } 0. \end{cases}$$

Notice that the quotient "$\dfrac{\varepsilon(x)}{\varepsilon(x)}$" does not make sense at all, so this notation should be avoided.

In some books $o(x)$ is used instead.
A function $f(x)$ is called an o-function, if $\dfrac{f(x)}{x}$, $x \neq 0$, is an ε-function. Applications with o-functions are more elegant, but may lead to misunderstandings for beginners. Therefore, we shall here only consider ε-functions.

It follows from the definition above that

$$f(x) = \varepsilon\,(x - x_0)\,, \qquad \text{if } f(x) \to 0 \text{ for } x \to x_0,$$

and

$$f(x) = \varepsilon\left(\frac{1}{x}\right), \qquad \text{if } f(x) \to 0 \text{ for } x \to +\infty, \quad (\text{or } x \to -\infty).$$

In the latter case we may specify, if $x > 0$ or $x < 0$.

6.3 Taylor's formula.

We have already in Section 4.3 shown a variant of this example of a useful application of *partial integration*, so we shall only give the result here, cf. also formula (100).

Theorem 6.1 *Let I be an open interval and let $a \in I$ be a fixed point. If $f \in C^n(I)$ for some $n \in \mathbb{N}$, then*

(167)
$$f(x) = f(a) + \frac{f'(a)}{1!}\,(x-a) + \frac{f''(a)}{2!}\,(x-a)^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}\,(x-a)^{n-1} + \frac{(-1)^n}{(n-1)!}\int_x^a (t-x)^{n-1} f^{(n)}(t)\,dt$$

for every $x \in I$.

As shown in Section 4.3, the proof of (167) is only using successive partial integrations. In the more traditional proofs in Calculus one uses also the *mean value theorem*. If so, we get under the same assumptions as in Theorem 6.1 the *apparently* stronger result

$$(168)\quad f(x) = f(a) + \frac{f'(a)}{1!}\,(x-a) + \frac{f''(a)}{2!}\,(x-a)^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}\,(x-a)^{n-1} + \frac{1}{n!}\,f^{(n)}(\xi)\cdot(x-a)^n,$$

for every $x \in I$, where $\xi = \xi(x)$ is some specific point (usually hard to find explicitly; we only know that such a point exists) between a and x.

Since $\xi = \xi(x)$ usually cannot be found in a reasonable way, one must estimate the error term. If we let $I_{a,x} \subset I$ denote the interval between a and x, i.e.

$$I_{a,x} = [a, x] \text{ for } a < x, \qquad \text{and} \qquad I_{a,x} = [x, a] \text{ for } x < a.$$

then we have the error estimate of (168),

$$|R_{n-1}(\xi, x, a)| = \left|\frac{1}{n!}\,f^{(n)}(\xi)\right| \cdot |x-a| \le \frac{1}{n!}\,\max_{\xi \in I_{a,x}} \left|f^{(n)}(\xi)\right| \cdot |x-a|^n.$$

This should be compared with the error estimate of (167), where

$$\begin{aligned} |R_{n-1}(x,a)| &= \left|\frac{(-1)^n}{(n-1)!}\int_x^a (t-x)^{n-1} f^{(n)}(t)\,dt\right| \le \frac{1}{(n-1)!}\,\max_{t \in I_{a,x}} \left|f^{(n)}(t)\right| \cdot \left|\int_x^a (t-x)^{n-1} dt\right| \\ &= \frac{1}{n!}\,\max_{t \in I_{a,x}} \left|f^{(n)}(t)\right| \cdot |x-a|^n, \end{aligned}$$

i.e. we get precisely the same error estimate of (167) and (168). One may therefore choose, whether one will continue working with the explicit error formula

$$R_{n-1}(x,a) := \frac{(-1)^n}{(n-1)!}\int_x^a (t-x)^{n-1} f^{(n)}(t)\,dt,$$

or with the more abstract error formula

$$R_{n-1}(\xi,x,a) := \frac{1}{n!} f^{(n)}(\xi)\cdot(x-a)^n,$$

where $\xi = \xi(x)$ is some point between a and x, which cannot be explicitly found.

I must say for one that I prefer the derivation of (167) to the derivation of (168), but when it comes to applications, (168) is in most cases the easiest one to apply, though I have proved above that they should always give the same error estimates. Notice also that we from (167) straight away get the following improvement of Theorem 6.1:

Since we have assumed that $f \in C^n$, we know that $f^{(n)}(x)$ is continuous, hence

$$f^{(n)}(x) - f^{(n)}(a) \to 0 \qquad \text{for } x \to a.$$

Then

$$\begin{aligned}
R_{n-1}(x,a) &= \frac{(-1)^n}{(n-1)!}\int_x^a (t-x)^{n-1} f^{(n)}(t)\,dt \\
&= \frac{(-1)^n}{(n-1)!} f^{(n)}(a)\int_x^a (t-x)^{n-1}\,dt + \frac{(-1)^n}{(n-1)!}\int_x^a (t-x)^{n-1}\left\{f^{(n)}(t) - f^{(n)}(a)\right\}dt \\
&= \frac{f^{(n)}(a)}{n!}\cdot(x-a)^n + \frac{(-1)^n}{(n-1)!}\int_x^a (t-x)^{n-1}\left\{f^{(n)}(t) - f^{(n)}(a)\right\}dt \\
&= \frac{f^{(n)}(a)}{n!}\cdot(x-a)^n + R_n(x,a),
\end{aligned}$$

where

$$(169)\quad R_n(x,a) := \frac{(-1)^n}{(n-1)!}\int_x^a (t-x)^{n-1}\left\{f^{(n)}(t) - f^{(n)}(a)\right\}dt = \frac{(x-a)^n}{n!}\cdot\varepsilon(x-a).$$

This shows that (167) for $f \in C^n$ can be written in the more natural way

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x,a),$$

where $R_n(x,a)$ is given by (169). We could also have obtained such a formula from (168), but then the error term (169) could only be given in the form $(x-a)^n\varepsilon(x-a)$, which however is sufficient in most cases.

From the considerations above it is seen that it is natural to introduce

Definition 6.1 *Let I be an open interval and $a \in I$ a given point. The* Taylor polynomial of order n *of a function $f \in C^n(I)$, expanded from $a \in I$ is defined by*

$$(170)\quad P_n(f,x,a) = P_n(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Remark 6.1 Notice that we write *order* n and not *degree* n in Definition 6.1. The reason is that if $f^{(n)}(a) = 0$, then the Taylor polynomial $P_n(x)$ of order n has degree $< n$. $\Diamond$

We can now give a better statement than Theorem 6.1:

Theorem 6.2 *Let I be an open interval, and let $a \in I$ be a point in I. If $f \in C^n(I)$, then*

$$\begin{aligned} f(x) &= f(a) + \frac{f'(a)}{1!}(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x,a) \\ &= P_n(x) + (x-a)^n \varepsilon(x-a), \end{aligned} \tag{171}$$

where $R_n(x,a)$ is given by (169).

It follows by a rearrangement that we have in particular

$$\frac{f(x) - P_n(x)}{(x-a)^n} = \varepsilon(x-a) \to 0 \qquad \text{for } x \to a. \tag{172}$$

Thus, the difference between $f \in C^n(I)$ and its Taylor polynomial, expanded from $a \in I$ can be divided by $(x-a)^n$, and the result will still tend to 0 for $x \to a$.

A more precise estimate of (172) is

$$(173)\quad \left|\frac{f(x)-P_n(x)}{(x-a)^n}\right| \leq \frac{1}{n!}\cdot \max_{t\in I_{x,a}} \left|f^{(n)}(t)-f^{(n)}(a)\right|,$$

which may be useful in some special cases.

6.4 Taylor expansions of standard functions.

Taylor's formula is valid for any function $f \in C^n(I)$ expanded from an interior point $a \in I$, but since our standard functions from Chapter 2 expanded from $a = 0$ occurs more frequently in the applications than most other functions, it is quite natural to collect these important Taylor expansions in one section.

All standard functions are of class C^∞ in their open domain, so if we let $a = 0$ and replace n with $n+1$ we have the following three variants of Taylor's formula

$$(174)\quad f(x) = f(0) + \frac{f'(0)}{1!}\,x + \frac{f''(0)}{2!}\,x^2 + \cdots + \frac{f^{(n)}(0)}{n!}\,x^n + \frac{(-1)^n}{n!}\int_0^x (t-x)^n f^{(n+1)}(t)\,dt,$$

$$(175)\quad f(x) = f(0) + \frac{f'(0)}{1!}\,x + \frac{f''(0)}{2!}\,x^2 + \cdots + \frac{f^{(n)}(0)}{n!}\,x^n + \frac{1}{(n+1)!}\,f^{(n+1)}(\xi)\cdot x^n,$$

$$(176)\quad f(x) = f(0) + \frac{f'(0)}{1!}\,x + \frac{f''(0)}{2!}\,x^2 + \cdots + \frac{f^{(n)}(0)}{n!}\,x^n + x^n\,\varepsilon(x),$$

where we only know that there exists a point $\xi \in I_{x,a}$, such that (175) holds.

Notice in (174) that

$$\frac{(-1)^{n+1}}{n!}\int_x^0 \cdots dt = \frac{(-1)^n}{n!}\int_0^x \cdots dt,$$

which explains the change of sign in the error term.

We see that we have three notations for the error term,

$$(177)\quad R_n(x,a) = \frac{(-1)^n}{n!}\int_0^x (t-x)^n f^{(n+1)}(t)\,dt = \frac{1}{(n+1)!}\,f^{(n+1)}(\xi)\,x^n = x^n\varepsilon(x),$$

which even might be supplied with the more complicated (169), if one likes this better. (I do not in this case, so avoid it.) The first one is the most precise, but difficult to handle in practice. The only reasonable case is when $f(x) = e^x$, so we shall only use it for the exponential.

The second one is easier to use in the applications, so we shall stick to that, although the students in general do not like this version, because $\xi \in I_{x,a}$ in most cases cannot be found with a reasonable small effort. However, we can always get rid of ξ by the following estimate,

$$(178)\quad |R_n(x)| = \frac{1}{(n+1)!}\left|f^{(n)}(\xi(x))\right|\cdot|x^n| \leq \max_{t\in I_{0,x}}\left|f^{(n)}(t)\right|\cdot\frac{|x|^n}{(n+1)!},$$

where we have changed the dubious ξ to t in the estimate to signal that ξ has left the problem.

Finally, the last notation, $R_n(x) = x^n \varepsilon(x)$ is more vague, concerning error estimates. It is used mainly, when we do *not* need a precise error estimate. This will be quite convenient when we later consider limits for $x \to 0$.

In applications in Physics one usually only considers the first two or three terms in an expansion. Therefore, an ideal scheme would be in each case to go through the following items,

1) Calculate the derivative $f^{(jn)}(x)$, $n \in \mathbb{N}$.

2) Specify the Taylor polynomial $P_n(x)$ in $f(x) = P_n(x) + R_n(x)$.

3) Give an estimate of the error term $R_n(x)$.

4) Give the usual approximations in Physics, supplied with graphs for comparison.

This scheme is not possible for all of our known functions without using a lot of unnecessary effort, so for some of the standard functions, in particular the inverse trigonometric functions and inverse hyperbolic functions, we shall give a somewhat simpler treatment.

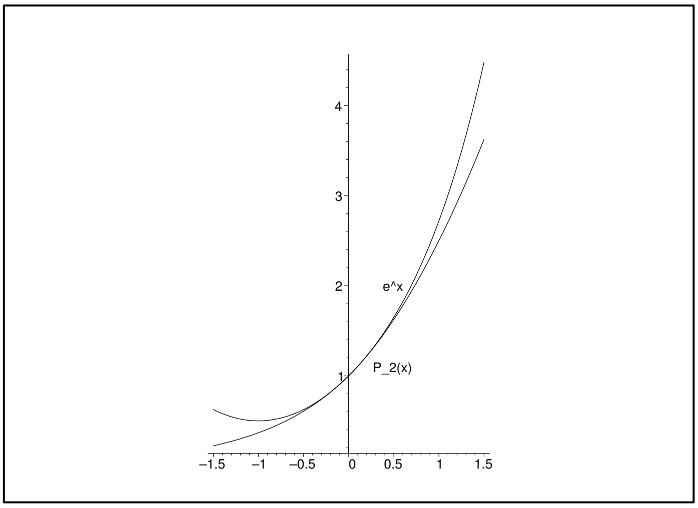

Figure 26: The graphs of e^x and $1 + x + \dfrac{x^2}{2}$.

The exponential e^x for $x \in \mathbb{R}$.

1) It is well-known that

$$f^{(n)}(x) = e^x \qquad \text{for all } n \in \mathbb{N}_0.$$

2) The Taylor expansion is

$$(179)\quad e^x = \exp x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + R_n(x).$$

3) The remainder term is according to (177) given by

$$(180)\quad R_n(x) = \frac{(-1)^n}{n!}\int_0^x (t-x)^n e^t\,dt = \frac{1}{(n+1)!}\,e^{\xi(x)}\cdot x^n = x^n\,\varepsilon(x).$$

Since e^x is increasing, we have the following error estimate for fixed $n \in \mathbb{N}$,

$$|R_n(x)| \le \begin{cases} \dfrac{e^x \cdot x^n}{(n+1)!} & \text{for } x > 0, \\[2ex] \dfrac{|x|^n}{(n+1)!} & \text{for } x < 0. \end{cases}$$

4) For $n = 2$ we get the simple approximation

$$e^x = 1 + x + \frac{1}{2}\,x^2 + x^2\,\varepsilon.$$

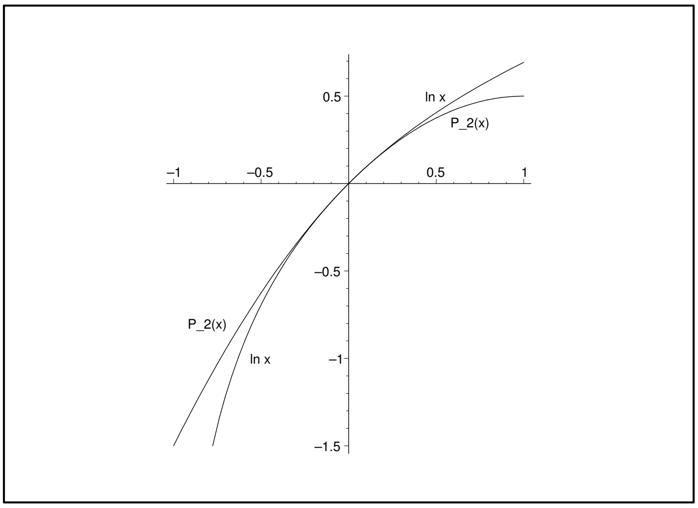

Figure 27: The graphs of $\ln(1+x)$ and $x - \frac{1}{2}\,x^2$.

The logarithm $\ln(1+x)$ **for** $x > -1$**.** Since we expand from $a = 0$, it is custom to consider $\ln(1+x)$ instead of $\ln x$, which is not defined for $x = 0$.

1) It is easily seen that

$$f^{(n)}(x) = (-1)^{n-1} \cdot \frac{(n-1)!}{(1+x)^n}, \qquad \text{for } n \in \mathbb{N}.$$

2) The Taylor expansion becomes

$$(181)\quad \ln(1+x) = x - \frac{x}{2} + \frac{x^2}{3} + \cdots + (-1)^{n-1}\,\frac{x^n}{n} + R_n(x), \qquad \text{for } x > -1.$$

3) The remainder term is here

$$(182)\quad R_n(x) = \frac{(-1)^n}{n+1} \cdot \frac{x^n}{(1+\xi)^{n+1}} = x^n\,\varepsilon(x),$$

where the exact expression of the error term in (177) already has become too complicated in order to be used in practice.

4) When $n = 2$, we get

$$\ln(1+x) = x - \frac{1}{2}\,x^{+}x^2\,\varepsilon(x).$$

The power function $(1+x)^\alpha$**,** $x > -1$ **and** $\alpha \in \mathbb{R}$**.**

1) It follows easily that

$$f^{(n)}(x) = \alpha(\alpha-1)\cdots(\alpha-n+1)\cdot(1+x)^{\alpha-1}.$$

2) The Taylor expansion becomes

$$(183)\quad (1+x)^{\alpha} = 1 + \frac{\alpha}{1!}\,x + \frac{\alpha(\alpha-1)}{2!}\,x^2 + \cdots + \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}\,x^n + R_n(x),$$

If we define the *generalized binomial coefficients* by

$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}, \qquad \alpha \in \mathbb{R}, \qquad n \in \mathbb{N}_0,$$

with n factors in both the numerator and the denominator, then (183) can be written

$$(184)\quad (1+x)^{\alpha} = 1 + \binom{\alpha}{1} x + \binom{\alpha}{2} x^2 + \cdots + \binom{\alpha}{n} x^n + R_n(x),$$

in agreement with the usual binomial formula, when $\alpha = n \in \mathbb{N}$. In the latter case we always get $R_n(x) \equiv 0$.

3) The remainder term is

$$(185)\quad R_n(x) = \binom{\alpha}{n+1}(1+\xi^{\alpha-n-1}\,x^n = x^n\,\varepsilon(x).$$

4) When $n = 2$ we get

$$(1+x)^{\alpha} = 1 + \alpha\,x + \frac{\alpha(\alpha-1)}{2}\,x^2 + x^2\,\varepsilon(x).$$

Concerning a figure we refer to the special case $\alpha = \frac{1}{2}$ in the following.

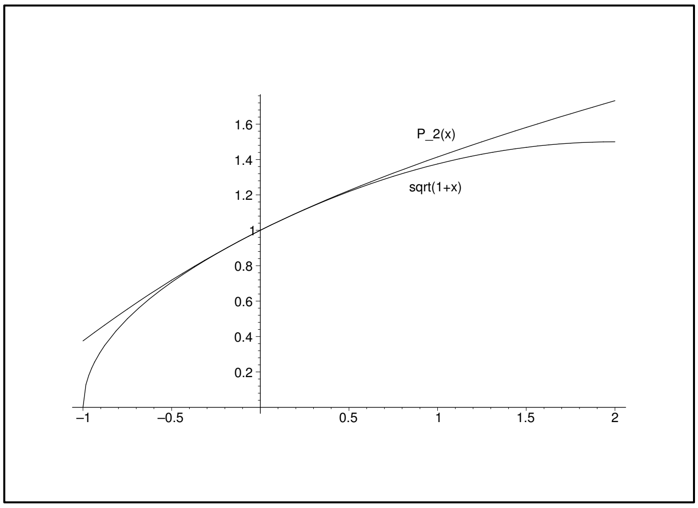

Figure 28: The graphs of $\sqrt{1+x}$ and $1+\frac{1}{2}\,x-\frac{1}{8}\,x^2$.

The square root $\sqrt{1+x}$ for $x > -1$. This is a special case, $\alpha = \frac{1}{2}$, of the former case.

1) We get after some computation and reduction,

$$f^{(n)}(x) = (-1)^{n-1}\frac{1}{2^n} \cdot 1 \cdot 3 \cdots (2n-3) \cdot \frac{\sqrt{1+x}}{(1+x)^n}, \qquad x > -1.$$

2) By some computation we get

$$\text{(186)} \quad \sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + \cdots + (-1)^{n-1} \cdot \frac{1}{2^{n-1}n} \binom{2n-2}{n-1} x^n + R_n(x).$$

3) The remainder term is here

$$\text{(187)} \quad R_n(x) = \frac{(-1)^n}{2(n+1)} \binom{2n}{n} \frac{\sqrt{1+\xi}}{(1+\xi)^{n+1}} \cdot x^n = x^n\,\varepsilon(x).$$

4) Choosing $n = 2$ we get the frequently used approximation in Physics,

$$\text{(188)} \quad \sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + x^2\,\varepsilon(x).$$

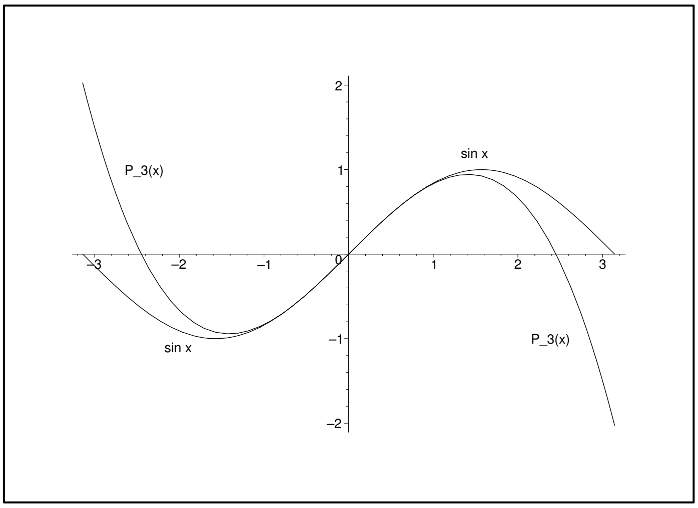

Figure 29: The graphs of $\sin x$ and $x - x^3/6$.

The trigonometric function $\sin x$, $x \in \mathbb{R}$.

1) We get

$$\begin{cases} f^{(2n)}(x) &= (-1)^n \sin x, \\ f^{(2n+1)}(x) &= (-1)^n \cos x, \end{cases} \qquad n \in \mathbb{N}_0.$$

2) Since $f^{(2n)}(0) = 0$ for all $n \in \mathbb{N}_0$, it is natural to find the Taylor expansion of order $2n+1$,

$$(189)\quad \sin x = \frac{x}{1!} - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + (-1)^n \cdot \frac{x^{2n+1}}{(2n+1)!} + R_{2n+1}(x).$$

3) The remainder term is

$$(190)\quad R_{2n+1}(x) = (-1)^n \cdot \frac{x^{2n+1}}{(2n+3)!}\,\sin\xi = x^{2n+2}\,\varepsilon(x).$$

4) For $n = 3$ we get

$$\sin x = x - \frac{x^3}{6} + x^4\,\varepsilon(x).$$

The trigonometric function $\cos x$, $x \in \mathbb{R}$.

1) We get

$$\begin{cases} f^{(2n)}(x) &= (-1)^n \cos x, \\ f^{(2n+1)}(x) &= (-1)^{n+1} \sin x, \end{cases} \qquad n \in \mathbb{N}_0.$$

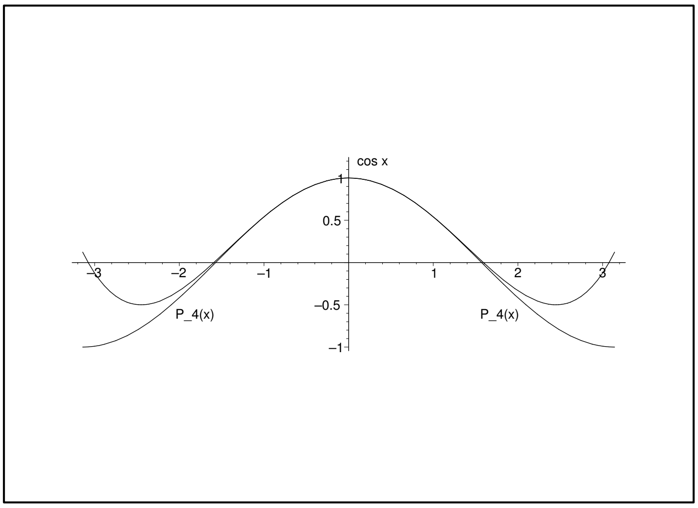

Figure 30: The graphs of $\cos x$ and $1 - \dfrac{1}{2}\,x^2 + \dfrac{1}{24}\,x^4$.

2) Since $f^{(2n+1)}(0) = 0$ for all $n \in \mathbb{N}_0$, it is natural to find the Taylor expansion of order $2n$,

(191) $\cos x = 1 - \dfrac{x^2}{2!} + \dfrac{x^4}{4!} + \cdots + (-1)^n\,\dfrac{x^{2n}}{(2n)!} + R_{2n}(x).$

3) The remainder term is

(192) $R_{2n}(x) = \dfrac{(-1)^{n+1}}{(2n+1)!}\,\sin\xi \cdot x^{2n+1} = x^{2n+1}\,\varepsilon(x).$

4) When $n = 4$, we get

$$\cos x = 1 - \frac{x^2}{2} + \frac{x^4}{24} + x^5\,\varepsilon(x).$$

The trigonometric functions $\tan x$ **and** $\cot x$**.** Only $\tan x$ has a Taylor expansion for $|x| > \frac{\pi}{2}$. Furthermore, it does not belong to a first year course in Calculus, so it is only given here for reference, when $n = 7$,

$$\tan x = x + \frac{1}{2}\,x^3 + \frac{2}{15}\,x^5 + \frac{17}{315}\,x^7 + x^8\,\varepsilon(x), \qquad \text{for } -\frac{\pi}{2} < x < \frac{\pi}{2}.$$

Since $\cot x$ is not defined for $x = 0$, it does not have a Taylor expansion. However, one can prove that $\cot x - \dfrac{1}{x}$ dose. We mention for completeness only that

(193) $$\cot x = \frac{1}{x} - \frac{1}{3}\,x - \frac{1}{45}\,x^3 - \frac{2}{945}\,x^5 + x^6\,\varepsilon(x), \qquad \text{for } 0 < |x| < \pi,$$

since this formula is not usually given in an elementary course in Calculus.

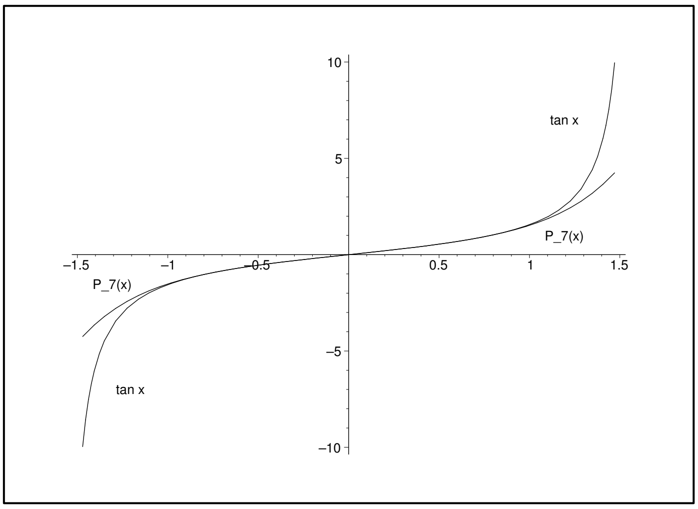

Figure 31: The graphs of $\tan x$ and $x + \dfrac{1}{3}\,x^3 + \dfrac{1}{15}\,x^5 + \dfrac{17}{315}\,x^7$ (not to scale).

Inverse trigonometric functions. When these were introduced in Chapter 2 it was proved that

(194) $$\operatorname{Arcsin} x + \operatorname{Arccos} x = \frac{\pi}{2} \qquad \text{and} \qquad \operatorname{Arctan} x + \operatorname{Arccot} x = \frac{\pi}{2}.$$

It is therefore sufficient only to give Taylor expansions of Arcsin x and Arctan x. Of these two functions, only the expansion of Arctan x is given in a first course in Calculus, because it is so important, and because it is easy to derive,

(195) $$\operatorname{Arctan} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots + (-1)^n\,\frac{x^{2n+1}}{2n+1} + x^{2n+2}\,\varepsilon(x).$$

It can be shown that one may use the following formula, when $x \in [-1,1]$,

(196) $\operatorname{Arctan} x = \dfrac{x}{1+0.28\,x^2} + R(x), \qquad x \in [-1,1],$

where

(197) $|R(x)| \leq \dfrac{1}{200} \qquad \text{for } x \in [-1,1].$

The Taylor expansion of Arcsin x is also considered as too advanced for a first course in Calculus. It is therefor here only mentioned for reference,

$$\operatorname{Arcsin} x = x + \frac{1}{6}\,x^3 + \frac{3}{40},x^5 + \cdots + \frac{1}{2^{2n}}\begin{pmatrix} 2n \\ n \end{pmatrix} \cdot \frac{1}{2n+1}\,x^{2n+1} + x^{2n+2}\,\varepsilon(x).$$

The hyperbolic functions $\sinh x$ **and** $\cosh x$**.** Since

$$\sinh x = \frac{1}{2}\left\{e^x - e^{-x}\right\} \qquad \text{and} \qquad \cosh x = \frac{1}{2}\left\{e^x + e^{-x}\right\},$$

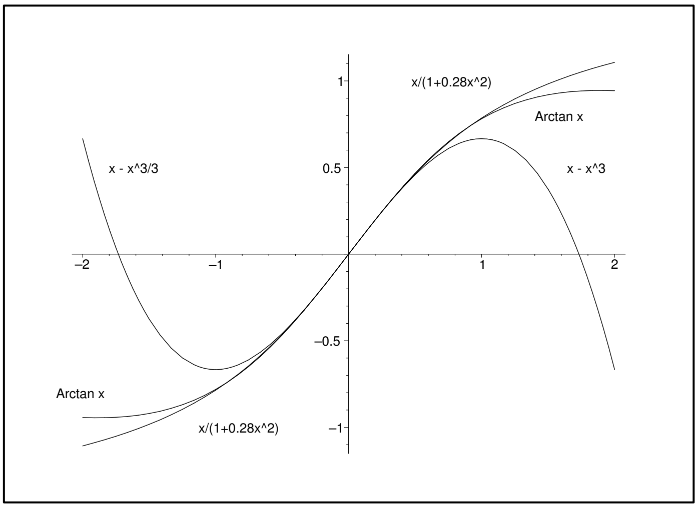

Figure 32: The graphs of Arctan x and $x - \frac{1}{3}\,x^3$ and $\frac{x}{1+0.28\,x^2}$ (not to scale).

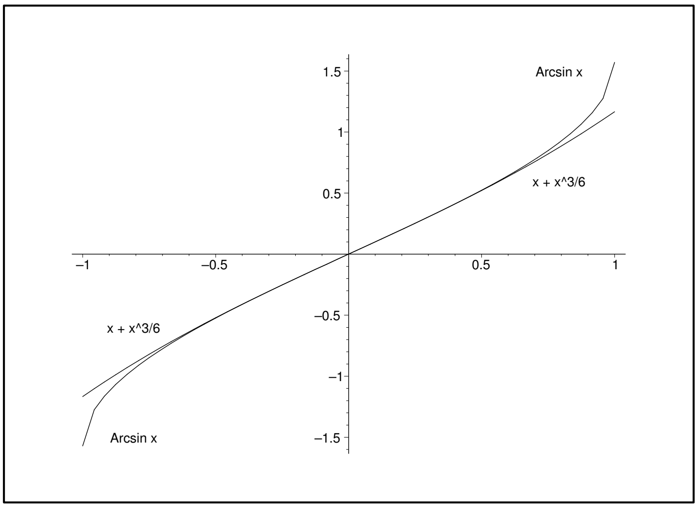

Figure 33: The graphs of Arcsin x and $x + \frac{1}{6}\,x^3$.

and the Taylor expansions are known for e^x and for e^{-x},

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^{2n}}{(2n)!} + \frac{x^{2n+1}}{(2n+1)!} + x^{2n+1}\,\varepsilon(x),$$

$$e^{-x} = 1 - \frac{x}{1!} + \frac{x^2}{2!} - \cdots + \frac{x^{2n}}{(2n)!} - \frac{x^{2n+1}}{(2n+1)!} + x^{2n+1}\,\varepsilon(x),$$

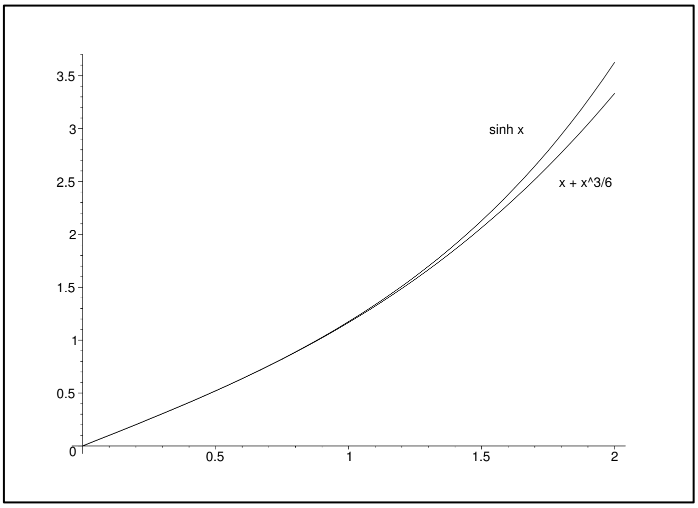

Figure 34: The graphs of $\sinh x$ and $x + \dfrac{1}{6}\,x^3$ (not to scale).

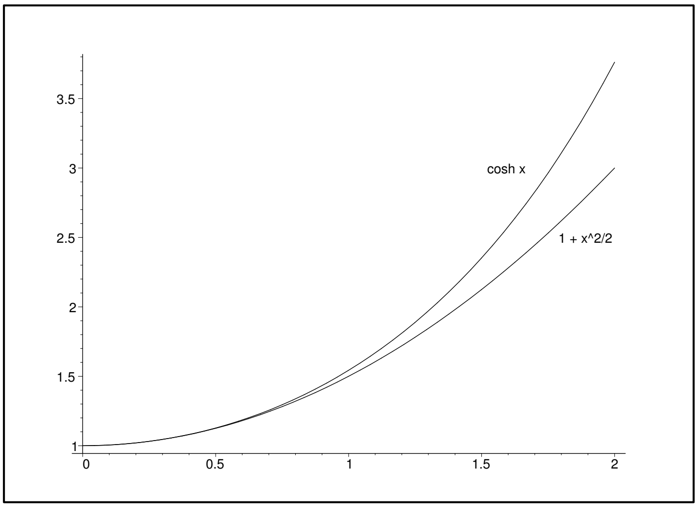

Figure 35: The graphs of $\cosh x$ and $1 + \dfrac{1}{2}\,x^2$ (not to scale).

it is easily seen that

$$(198)\quad \sinh x = \frac{x}{1!} + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + \frac{x^{2n+1}}{(2n+1)!} + x^{2n+1}\,\varepsilon(x),$$

$$(199)\quad \cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + \frac{x^{2n}}{(2n)!} + x^{2n}\,\varepsilon(x).$$

The hyperbolic functions $\tanh x$ **and** $\coth x$**.** these are usually not considered either in a first course of Calculus. We mention for reference only that

$$\tanh x = x - \frac{1}{3}\,x^3 + \frac{2}{15}\,x^5 - \frac{17}{315}\,x^7 + x^8\,\varepsilon(x), \qquad \text{for } -\frac{\pi}{2} < x < \frac{\pi}{2}.$$

The inverse hyperbolic functions. These are usually not included in a first course in Calculus. Only Arsinh x and Artanh x have reasonable expansions, which here only are mentioned for reference,

(200)

$$\operatorname{Arsinh} x = x - \frac{1}{2\cdot 3}\,x^3 + \frac{1\cdot 3}{2\cdot 4\cdot 5}\,x^5 + \cdots + (-1)^n \cdot \frac{1}{2^n}\binom{2n}{n}\frac{x^{2n+1}}{2n+1} + x^{2n+2}\,\varepsilon(x), \qquad \text{for } -1 < x < 1,$$

and

$$\operatorname{Artanh} x = x + \frac{x^3}{3} + \frac{x^5}{5} + \cdots + \frac{x^{2n+1}}{2n+1} + x^{2n+2}\,\varepsilon(x), \qquad \text{for } -1 < x < 1.$$

The strange phenomenon that (200) can only be used for $|x| < 1$, although Arsinh x is defined for all $x \in \mathbb{R}$, cannot be explained without some knowledge of e.g. Complex Function Theory.

6.5 Limits.

We shall next consider the problem of finding the limit

$$(201) \quad \lim_{x \to a} \frac{f(x)}{g(x)},$$

where both $f(x)$ and $g(x)$ are of class C^n or C^∞ in a neighbourhood of the point a. We shall here give a procedure of solving this problem by using Taylor expansions of the functions.

1) Always first consider the denominator $g(x)$. If it explicitly includes a factor of the type $(x-a)^p$, $p \in \mathbb{N}$, we write

$$g(x) = (x-a)^p g_1(x),$$

where $p \in \mathbb{N}_0$ is the highest exponent for which this is possible. If no such factor exist, we have $p = 0$ and $g(x) = g_1(x)$.

2) Compute

$$g_1(a), \qquad g_1'(a), \qquad g_1''(a), \qquad \dots,$$

until we for the first time obtain $g_1^{(q)}(a) \neq 0$. Then by Taylor's formula,

$$g(x) = \frac{1}{q!}\,(x-a)^{p+q} g_1 * (q)(a) + (x-a)^{p+q}\,\varepsilon(x-a),$$

and $p+q$ is called the *order of the zero at $x = a$ of the denominator.*

3) Expand $f(x)$ in the same way *of the same order* $p+q$, which has been defined by the expansion of $g(x)$ above, i.e. first we write as in 1)

$$f(x) = (x-a)^r f_1(x), \qquad r \in \mathbb{N}_0,$$

and then calculate

$$f_1(a), \qquad f_1'(a), \qquad f_1''(a), \qquad \dots, \qquad f_1^{(p+q-r)}(a).$$

4) If $f_1^{(j)}(a) \neq 0$ for some $j < p+q-r$, then the limit (201) does not exist.

5) If $f_1^{(j)}(a) = 0$ for all $j < p+q-r$, we get

$$f(x) = \frac{f_1^{(p+q-r)}(a)}{(p+q-r)!}\,(x-a)^{p+q} + (x-1)^{p+q}\,\varepsilon(x-a),$$

and hence by insertion in (201),

$$\begin{aligned}
\frac{f(x)}{g(x)} &= \frac{q!}{(p+q-r)!} \cdot \frac{f_1^{(p+q-r)}(a) \cdot (x-a)^{p+q} + (x-a)^{p+q}\,\varepsilon(x-a)}{g_1^{(q)}(a) \cdot (x-a)^{p+q} + (x-a)^{p+q}\,\varepsilon(x-a)} \\
&= \frac{q!}{(p+q-r)!} \cdot \frac{f_1^{(p+q-r)}(a) + \varepsilon(x-a)}{g_1^{(q)}(a) + \varepsilon(x-a)} \\
&\to \frac{q!}{(p+q-r)!} \cdot \frac{f_1^{(p+q-r)}(a)}{g_1^{(q)}(a)} \qquad \text{for } x \to a,
\end{aligned}$$

and we have found the limit of (201).

Remark 6.2 It is extremely important that we find the first $g_1^{(q)}(a) \neq 0$ in the denominator. Therefore, one should always start by finding the order of the zero of the denominator in $x = a$, and let this order define the order of expansion of the numerator. $\Diamond$

Example 6.1 The classical (and easy) example is the limit of

$$\frac{f(x)}{g(x)} = \frac{\sin x}{x} \qquad \text{for } x \to 0.$$

The denominator $g(x) = x = (x-0)^1$ has already the right structure, where the order is 1 at $x = 0$, so we get for the numerator,

$$f(x) = \sin x = x + x^2\,\varepsilon(x).$$

By insertion we get

$$\frac{\sin x}{x} = \frac{x + x^2\,\varepsilon(x)}{x} = \frac{1 + x\,\varepsilon(x)}{1} \to 1 \qquad \text{for } x \to 0,$$

so we have derived the well-known result

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

Example 6.2 Consider

$$\frac{\cot x}{x} - \frac{1}{x^2} = \frac{x\,\cos x - \sin x}{x^2\,\sin x} = \frac{f(x)}{g(x)}, \qquad x \neq p\pi, \quad p \in \mathbb{Z},$$

for $x \to 0$.

It follows immediately that

$$g(x) = x^2\,\sin x = x^2\,g_1(x), \quad \text{where} \quad g_1(x) = \sin x = x + x^2\,\varepsilon(x),$$

so the denominator can be written

$$g(x) = x^3 + x^4\,\varepsilon(x) = x^3\,\{1 + x\,\varepsilon(x)\},$$

i.e. it has a zero of order 3 at $x = 0$.

when the numerator $f(x)$ is expanded of order 3, we get

$$\begin{aligned}
f(x) &= x \cdot \cos x - \sin x \\
&= x\left\{1 - \frac{1}{2}\,x^2 + x^3\,\varepsilon(x)\right\} - \left\{x - \frac{1}{6}\,x^3 + x^4\,\varepsilon(x)\right\} \\
&= \left\{x - \frac{1}{2}\,x^3 + x^4\,\varepsilon(x)\right\} - \left\{x - \frac{1}{6}\,x^3 + x^4\,\varepsilon(x)\right\} \\
&= -\frac{1}{3}\,x^3 + x^4\,\varepsilon(x) = x^3\left\{-\frac{1}{3} + x\,\varepsilon(x)\right\}.
\end{aligned}$$

By insertion we get

$$\begin{aligned}
\frac{\cot x}{x} - \frac{1}{x^2} &= \frac{x \cdot \cos x - \sin x}{x^2\,\sin x} = \frac{f(x)}{g(x)} = \frac{x^3\left\{-\dfrac{1}{3} + x\,\varepsilon(x)\right\}}{x^3\,\{1 + x\,\varepsilon(x)\}} \\
&= \frac{-\dfrac{1}{3} + x\,\varepsilon(x)}{1 + x\,\varepsilon(x)} \to -\frac{1}{3} \qquad \text{for } x \to 0,
\end{aligned}$$

and it follows that

$$\lim_{x\to 0}\left\{\frac{\cot x}{x}-\frac{1}{x^2}\right\}=-\frac{1}{3}. \qquad \Diamond$$

6.6 Asymptotes.

In the investigation of a curve with the equation $y=f(x)$ we may in some cases be able to approximate the graph with a straight line, "when we are far away from $(0,0)$" in some sense. Such lines are called *asymptotes*.

1) **Vertical asymptotes**. A vertical asymptote occurs when

$$|f(x)|\to ++\infty, \qquad \text{for } x\to a+ \text{ or } x\to a- .$$

Typical examples of vertical asymptotes are $x=\dfrac{\pi}{2}+p\pi$, $p\in\mathbb{Z}$, for the function $y=\tan x$.

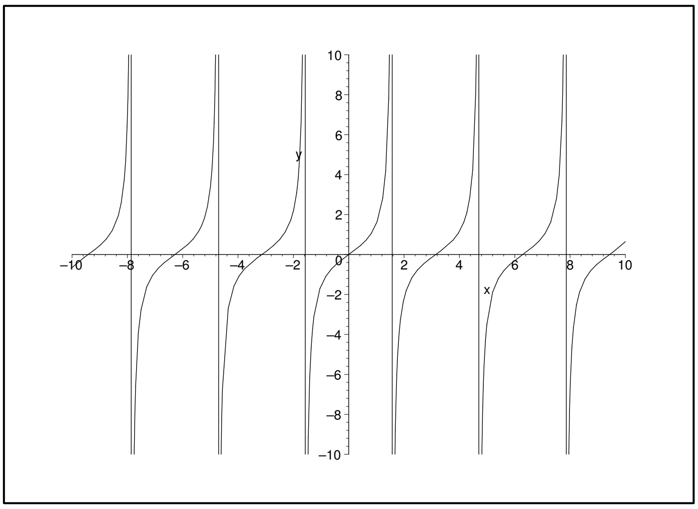

Figure 36: The graphs of $\tan x$ and some of its asymptotes.

2) **Asymptotes of the type $y=ax+b$.** Let $f(x)$ be C^1 in the interval $I=]\alpha,+\infty[$ (resp. $I=]-\infty,\alpha[$). We say that the graph of $y=f(x)$ has the line $y=ax+b$ as an *asymptote*, if

(202) $f(x)-\{ax+b\}\to 0 \qquad \text{for } x\to+\infty \qquad (x\to-\infty,\ \text{resp.}).$

If we somehow can write $f(x)$ in the form

(203) $f(x)=ax+b+\varepsilon\left(\dfrac{1}{x}\right), \qquad x\to+\infty \qquad (\text{or } x\to-\infty),$

then of course $y=ax+b$ is an asymptote for $y=f(x)$.

Example 6.3 The classical example is there the asymptote of the function (part of a hyperbola)

(204) $y=\sqrt{x^2+1}, \qquad x\geq 0.$

In this case we get for $x > 0$ by using the standard expansion of the square root,

$$\begin{aligned} y &= \sqrt{x^2+1} = x\sqrt{1+\frac{1}{x^2}} = x\left\{1+\frac{1}{2}\cdot\frac{1}{x^2}+\frac{1}{x^2}\,\varepsilon\left(\frac{1}{x}\right)\right\} \\ &= x+\frac{1}{x}+\frac{1}{x}\,\varepsilon\left(\frac{1}{x}\right) = x+\varepsilon\left(\frac{1}{x}\right) \qquad \text{for } x \to +\infty, \end{aligned}$$

so $y = x$ is an asymptote. $\Diamond$

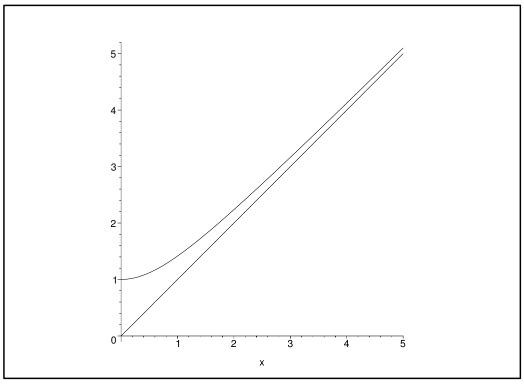

Figure 37: Graph of $y = \sqrt{x^2+1}$, $x \geq 0$ and its asymptote $y = x$.

Example 6.4 If instead $f(x) = \sqrt{x^2+x+1}$, $x \geq 0$, we bet by the same method as in Example 6.3 for $x > 0$,

$$\begin{aligned} f(x) &= \sqrt{x^2+x+1} = x\sqrt{1+\frac{1}{x}+\frac{1}{x^2}} = x\left\{1+\frac{1}{2}\left(\frac{1}{x}+\frac{1}{x^2}\right)+\frac{1}{x}\,\varepsilon\left(\frac{1}{x}\right)\right\} \\ &= x+\frac{1}{2}+\varepsilon\left(\frac{1}{x}\right), \end{aligned}$$

from which we conclude that $y = \sqrt{x^2+x+1}$ has the asymptote $y = x + \dfrac{1}{2}$ for $x \to +\infty$. $\Diamond$

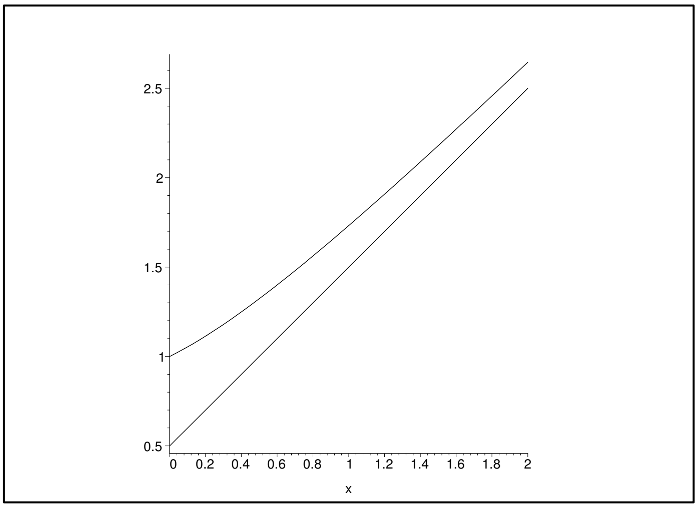

Figure 38: The function $y = \sqrt{x^2+x+1}$ and its asymptote $y = x + \dfrac{1}{2}$.

Example 6.5 *Examine the function*

$$f(x) = x\,\tanh\sqrt{x}$$

for a possible asymptote for $x \to +\infty$.

In this case we cannot use Taylor's formula, but we can instead use the rules of magnitude. By a rearrangement we get

$$\begin{aligned} x\cdot\tanh\sqrt{x} &= x\cdot\frac{\sinh\sqrt{x}}{\cosh\sqrt{x}} = x\cdot\frac{e^{\sqrt{x}}-e^{-\sqrt{x}}}{e^{\sqrt{x}+e^{-\sqrt{x}}}} = x\cdot\frac{1-e^{-2\sqrt{x}}}{1+e^{-2\sqrt{x}}} \\ &= x\cdot\left\{1-2\cdot\frac{e^{-2\sqrt{x}}}{1+e^{-2\sqrt{x}}}\right\} = x-2\cdot\frac{x\cdot e^{-2\sqrt{x}}}{1+e^{-2\sqrt{x}}}. \end{aligned}$$

If we put $u = \sqrt{x}$, we get

$$\left|-2\cdot\frac{x\cdot e^{-2\sqrt{x}}}{1+e^{-2\sqrt{x}}}\right| = \frac{2}{1+e^{-2u}}\cdot\frac{u^2}{e^{2u}} < 2\left(\frac{u}{e^u}\right)^2 \to 0 \qquad \text{for } u = \sqrt{x} \to +\infty,$$

where we have used that an exponential dominates a power function.

It follows that

$$x \cdot \tanh \sqrt{x} = x + \varepsilon \left(\frac{1}{x}\right),$$

so $y = x$ is an asymptote for $y = x \cdot \tanh \sqrt{x}$. $\diamondsuit$

Figure 39: The function $y = x \cdot \tanh \sqrt{x}$, $x \geq 0$, and its asymptote $y = x$.

6.7 Approximations of integrals.

It is possible to use Taylor's formula to find approximate values of integrals in intervals which are not too large. Although this is a standard procedure I am very reluctant to describe the method here. The reason is that one in most cases can find other methods which are much better than using a Taylor expansion.

In order to illustrate that a Taylor expansion is not always the most economical way of estimating an integral we shall consider the classical *error function* defined by

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt.$$

It is of course sufficient to calculate the integral

$$F(x) = \int_0^x e^{-t^2}\, dt$$

without the normalizing factor $\dfrac{2}{\sqrt{\pi}}$.

We proved in Chapter 4 by using partial integration successively that

$$F(x) = \int_0^x e^{-t^2}\,dt = e^{-x^2}\left\{x + \frac{2}{1\cdot 3}\,x^3 + \frac{2^2}{1\cdot 3\cdot 5}\,x^5\right\} + \frac{2^3}{1\cdot 3\cdot 5}\int_0^x t^6 e^{-t^2}\,dt.$$

It is easily proved by induction that we have more generally

$$\int_0^x e^{-t^2}\,dt = e^{-x^2}\left\{x + \frac{2}{1\cdot 3}\,x^3 + \cdots + \frac{2^n}{1\cdot 3\cdots(2n+1)}\,x^{2n+1}\right\} + \frac{2^{n+1}}{1\cdot 3\cdots(2n+1)}\int_0^x t^{2n+2}e^{-t^2}\,dt$$

for all $n \in \mathbb{N}$.

When $x > 0$ the remainder term is also > 0, so by a rough estimate,

$$\begin{aligned} 0 < R_{2n+1}(x) &= \frac{2^{n+1}}{1\cdot 3\cdots(2n+1)}\int_0^x t^{n+2}e^{-t^2}\,dt < \frac{2^{n+1}}{1\cdot 3\cdots(2n+1)}\int_0^x t^{2n+2}\,dt \\ &= \frac{2^{n+1}\cdot x^{2n+3}}{1\cdot 3\cdots(2n+1)(2n+3)}. \end{aligned}$$

Since a faculty function dominates a power function it is for any given $x > 0$ and $\varepsilon > 0$ possible to find $n_0 \in \mathbb{N}$, such that

$$0 < R_{2n+1}(x) < \frac{2^{n+1} \cdot x^{2n+3}}{1 \cdot 3 \cdots (2n+1)(2n+3)} < \varepsilon, \qquad \text{for all } n \geq n_0.$$

Hence, by using this method we can calculate $\operatorname{erf}(x)$ for any given $x > 0$ with as small error as we with.

For *comparison* we now calculate the same integral $\int_0^x e^{-t^2}\,dt$, $x > 0$, where we this time use Taylor's formula.

Calculate the first derivatives of the integrand $f(x) = e^{-x^2}$, i.e.

$$\begin{aligned}
f'(x) &= -2x\,e^{-x^2}, \\
f''(x) &= -2\,e^{-x^2} + 4x^2 e^{-x^2}, \\
f^{(3)}(x) &= 12x\,e^{-x^2} - 8x^3\,e^{-x^2}, \\
f^{(4)}(x) &= 12\,e^{-x^2} - 48x^2 e^{-x^2} + 16x^4 e^{-x^2} \\
f^{(5)}(x) &= -120x\,e^{-x^2} + 160x^3 e^{-x^2} - 32x^5 e^{-x^2}, \\
f^{(6)}(x) &= -120\,e^{-x^2} + 720x^2 e^{-x^2} - 48 + x^4 e^{-x^2} + 64x^6 e^{-x^2}, \\
f^{(7)}(x) &= 1680x\,e^{-x^2} - 3360x^3 e^{-x^2} + 1344x^5 e^{-x^2} - 128x^7 e^{-x^2},
\end{aligned}$$

(these calculations really grow wildly!) so

$$f(0) = 1, \quad f'(0) = 0, \quad f''(0) = -2, \quad f^{(3)}(0) = 0, \quad f^{(4)}(0) = 12, \quad f^{(5)}(0) = 0, \quad f^{(6)}(0) = 120,$$

and

$$f^{(7)}(x) = 16x\,e^{-x^2}\left\{105 - 210x^2 + 84x^4 - 8x^6\right\}.$$

It follows that

$$e^{-x^2} = 1 - \frac{2}{2!}\,x^2 + \frac{12}{4!}\,x^4 - \frac{120}{6!}\,x^6 + \frac{f^{(7)}(\xi)}{7!}\,x^7 = 1 - x^2 + \frac{1}{2}\,x^4 - \frac{1}{6}\,x^6 + \frac{f^{(7)}(\xi)}{5040}\,x^7.$$

We get by an integration,

$$\int_0^x e^{-t^2}\,dt = x - \frac{1}{3}\,x^3 + \frac{1}{10}\,x^5 - \frac{1}{42}\,x^7 + \int_0^x \frac{f^{(7)}(\xi(t))}{5040}\,t^7\,dt.$$

The usual estimate gives here

$$\left|f^{(7)}(\xi)\right| \leq 16 \cdot 1 \cdot 1 \cdot \{105 + 210 + 84 + 8\} = 16 \cdot 407 = 6512,$$

hence

$$\left|\int_0^x \frac{f^{(7)}(\xi(t))}{5040} \cdot t^7\,dt\right| \leq \frac{6512}{5040} \cdot \frac{1}{8} \cdot x^8 = \frac{407}{2520}\,x^8.$$

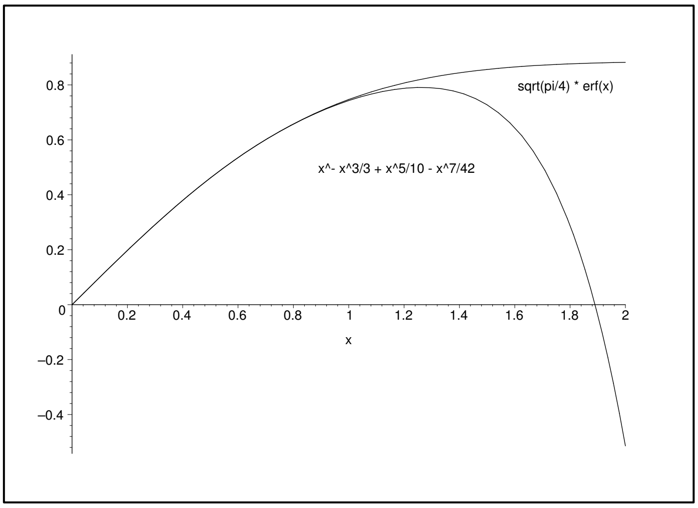

Figure 40: The graphs of $F(x) = \int_0^x e^{-t^2}\,dt$, $x \geq 0$, and its approximation $y = x - \frac{1}{3}\,x^3 + \frac{1}{10}\,x^5 - \frac{1}{42}\,x^7$.

This estimate is reasonable when e.g. $0 < x < \frac{1}{2}$, but when $x > 1$ it is not very good. It is seen from the figure that the polynomial does not give a good approximation for $x \geq 1.2$.

Remark 6.3 Until approximately 30 years ago one often saw exercises where one should give an estimate of $f^{(n)}(\xi)$ for ξ lying between a and x. These exercises were always feared by the student, because they were very difficult. Seen in retrospect these exercises did not add much, because there exist better approximation methods, and the vague notation $\varepsilon(x)$ was also more easy to handle, although one of course lost some information. These estimates of $f^{(n)}(\xi)$ were therefore rightfully abandoned from the common courses in Calculus, but one still cannot deny that thy could be useful in some very special cases. $\Diamond$

6.8 Miscellaneous applications.

Taylor expansions may be used to give explicit approximations for *implicitly given functions*. Although the theory does not belong to a first course in Calculus, it is not difficult already here to give some computational clues. We shall not here be mathematically precise, only illustrate the possible method by suitable examples.

Example 6.6 *Find an approximation Taylor polynomial ($a = 0$) for the implicitly given function*

(205) $x \cdot \ln x = t, \qquad x > 0.$

Notice that when $t = 0$, then $x = 1$, because we have assumed that $x > 0$, hence $x(0) = 1$.

It can be proved (Calculus 2) that (205) in a neighbourhood of $t = 0$ can be described as a C^∞-function. We shall here take this for granted.

By differentiation of (205), using that $\ln x(1) = \ln 1 = 0$, we get

(206) $(1 + \ln x)\,\dfrac{dx}{dt} = 1, \qquad \text{hence } x'(0) = 1.$

By another differentiation and a rearrangement we get

$$(1 + \ln x)\,\frac{d^2x}{dt^2} = -\frac{1}{x}\left(\frac{dx}{dt}\right)^2, \qquad \text{hence } x''(0) = -1.$$

Using the same procedure once again we obtain

$$(1 + \ln x)\,\frac{d^3x}{dt^3} = -\frac{3}{x}\cdot\frac{dx}{dt}\cdot\frac{d^2x}{dt^2} + \frac{1}{x^2}\left(\frac{dx}{dt}\right)^3, \qquad \text{hence } x^{(3)}(0) = 4.$$

Since

$$x(0) = 1, \qquad x'(0) = 1, \qquad x''(0) = -1 \quad \text{and} \quad x^{(3)}(0) = 4,$$

we get by Taylor's formula that

$$x(t) = 1 + t - \frac{1}{2}\,t^2 + \frac{2}{3}\,t^3 + t^3\,\varepsilon(t).$$

In this way we continue as long as we wish. $\Diamond$

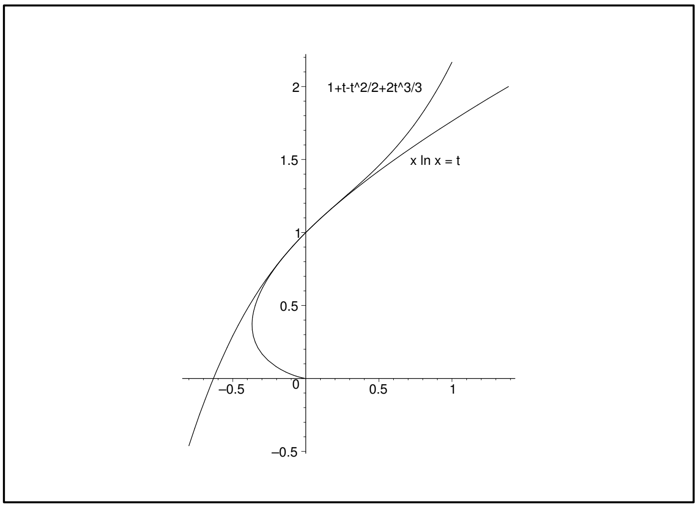

Figure 41: The graphs of $x \cdot \ln x = t$ and $x = 1 + t - \frac{1}{2}\,t^2 + \frac{2}{3}\,t^3$. Notice that the graph of the implicit given function bends back towards the Y-axis, so it cannot be described as a function in t in the whole range where it is defined.

Example 6.7 It is seen from Example 6.6 that the implicitly given function (205) can also be described by the differential equation (206), supplied with the initial condition $x(0) = 1$, hence

(207) $$(1 + \ln x)\,\frac{dx}{dt} = 1, \qquad x(0) = 1.$$

One obtains an approximate solution of (207) by differentiation precisely like in Example 6.6. Notice that when $t > -\dfrac{1}{e}$, then (207) is equivalent to

$$(208)\quad \frac{dx}{dt} = \frac{1}{1+\ln x}, \qquad x(0) = 1.$$

The method of differentiation can be applied on both (207) and (208), and the resulting approximating Taylor polynomial is of course the same, but the differentiations of (208) are much harder to perform than the differentiations of (207). One should therefore always choose the differential equation which is used with some care. $\Diamond$

Example 6.8 Equation (207) is a nonlinear differential equation, where the variables can be separated. This is not always the case for nonlinear equations.

As an example we return to the equation (164) in Section 6.1 of the *mathematical pendulum*, i.e.

$$(209)\quad \frac{d^2\varphi}{dt^2} = -c \cdot \sin\varphi, \qquad \text{where } c = \frac{g}{\ell} \text{ and } \varphi(0) = 0 \text{ and } \varphi'(0) = a.$$

No nontrivial exact solution of (209) is known, but it is still possible to find a Taylor approximation of the unique solution of (209).

By putting $t = 0$ in the equation (209) we get $\varphi''(0) = 0$, so

$$\varphi(0) = 0, \qquad \varphi'(0) = a \quad \text{and} \quad \varphi''(0) = 0.$$

When (209) is differentiated once again we get

$$\frac{d^3\varphi}{dt^3} = -c \cdot \cos\varphi \cdot \frac{d\varphi}{dt}, \qquad \text{hence } \varphi^{(3)}(0) = -ac.$$

In the next step we obtain

$$\frac{d^4\varphi}{dt^4} = -c \cdot \cos\varphi \cdot \frac{d^2\varphi}{dt^2} + c \cdot \sin\varphi \cdot \left(\frac{d\varphi}{dt}\right)^2, \qquad \text{hence } \varphi^{(4)}(0) = 0.$$

Finally,

$$dfracd^5\varphi dt^5 = -c \cdot \cos\varphi \cdot \frac{d^3\varphi}{dt^3} + 3c \cdot \sin\varphi \cdot \frac{d\varphi}{dt} \cdot \frac{d^2\varphi}{dt^2} + c \cdot \cos\varphi \cdot \left(\frac{d\varphi}{dt}\right)^3,$$

hence for $t = 0$,

$$\varphi^{(5)}(0) = -c \cdot \varphi^{(3)}(0) + 0 + c \cdot \{\varphi'(0)\}^3 = -c(-ac) + c \cdot a^3 = ac^2 + a^3 c.$$

Using that $c = \dfrac{g}{\ell}$ we get the Taylor approximation

$$(210)\quad \varphi(t) = at - \frac{ag}{3!\ell}\, t^3 + \frac{1}{5!}\left\{\frac{ag^2}{\ell^2} + \frac{a^3 g}{\ell}\right\} t^5 + t^5\,\varepsilon(t).$$

Notice that all terms have no physical dimension. The nonlinearity of the problem is obvious, e.g. by considering the structure of the coefficient of t^5. One obtains a better accuracy by repeating this method of finding the Taylor coefficients. $\Diamond$

Example 6.9 Let us consider the nonlinear equation

$$(211)\quad \frac{dx}{dt} = t + x^2, \qquad x(0) = 0,$$

(a Riccati equation). In most cases Riccati equations cannot be solved explicitly. There are, however, a few not well-known exceptions, and I cannot say whether (211) is one of them or not. We can, however, still use the method of differentiation to find an approximate solution of (211).

If follows immediately from (211) that $x'(0) = 0$. By a differentiation we get

$$\frac{d^2x}{dt^2} = 1 + 2x\,\frac{dx}{dt}, \qquad \text{hence } x''(0) = 1.$$

When this procedure is iterated we get

$$\frac{d^3x}{dt^3} = 2x\,\frac{d^2x}{dt^2} + 2\left(\frac{dx}{dt}\right)^2, \qquad \varphi^{(3)}(0) = 0,$$

$$\frac{d^4x}{dt^4} = 2x\,\frac{d^3x}{dt^3} + 6\,\frac{dx}{dt}\,\frac{d^2x}{dt^2}, \qquad \varphi^{(4)}(0) = 0,$$

$$\frac{d^5x}{dt^5} = 2x\,\frac{d^4x}{dt^4} + 8\,\frac{dx}{dt}\cdot\frac{d^3x}{dt^3} + 6\left(\frac{d^2x}{dt^2}\right)^2, \qquad \varphi^{(5)}(0) = 6,$$

etc.. At this stage the process the Taylor approximation of the solution of (211) is given by

$$(212)\quad x = \frac{1}{2}\,t^2 + \frac{1}{20}\,t^5 + t^5\,\varepsilon(t).$$

It is easily seen that $\varphi^{(8)}(0)$ is the next derivative $\neq 0$, so (212) may even be written

$$x = \frac{1}{2}\,t^2 + \frac{1}{20}\,t^5 + t^7\,\varepsilon(t). \qquad \Diamond$$

A Formulæ

Some of the following formulæ can be assumed to be known from high school. Others are introduced in Calculus 1. It is highly recommended that one *learns most of these formulæ in this appendix by heart.*

A.1 Squares etc.

The following simple formulæ occurs very frequently in the most different situations.

$$
\begin{array}{ll}
(a+b)^2 = a^2+b^2+2ab, & a^2+b^2+2ab = (a+b)^2, \\
(a-b)^2 = a^2+b^2-2ab, & a^2+b^2-2ab = (a-b)^2, \\
(a+b)(a-b) = a^2-b^2, & a^2-b^2 = (a+b)(a-b), \\
(a+b)^2 = (a-b)^2+4ab, & (a-b)^2 = (a+b)^2-4ab.
\end{array}
$$

A.2 Powers etc.

Logarithm:

$$
\begin{array}{rcll}
\ln|xy| &=& \ln|x|+\ln|y|, & x,y\neq 0, \\
\ln\left|\dfrac{x}{y}\right| &=& \ln|x|-\ln|y|, & x,y\neq 0, \\
\ln|x^r| &=& r\ln|x|, & x\neq 0.
\end{array}
$$

Power function, fixed exponent:

$$(xy)^r = x^r\cdot y^r, x,y>0 \quad \text{(extensions for some } r\text{)},$$

$$\left(\frac{x}{y}\right)^r = \frac{x^r}{y^r}, x,y>0 \quad \text{(extensions for some } r\text{)}.$$

Exponential, fixed base:

$$
\begin{array}{ll}
a^x\cdot a^y = a^{x+y}, \quad a>0 & \text{(extensions for some } x,\, y\text{)}, \\
(a^x)^y = a^{xy}, a>0 & \text{(extensions for some } x,\, y\text{)},
\end{array}
$$

$$a^{-x} = \frac{1}{a^x}, a>0, \qquad \text{(extensions for some } x\text{)},$$

$$\sqrt[n]{a} = a^{1/n}, a\geq 0, \qquad n\in\mathbb{N}.$$

Square root:

$$\sqrt{x^2} = |x|, \qquad x\in\mathbb{R}.$$

Remark A.1 It happens quite frequently that students make errors when they try to apply these rules. They must be mastered! In particular, as one of my friends once put it: “If you can master the square root, you can master everything in mathematics!” Notice that this innocent looking square root is one of the most difficult operations in Calculus. Do not forget the *absolute value!* $\Diamond$

A.3 Differentiation

Here are given the well-known rules of differentiation together with some rearrangements which sometimes may be easier to use:

$$\{f(x) \pm g(x)\}' = f'(x) \pm g'(x),$$

$$\{f(x)g(x)\}' = f'(x)g(x) + f(x)g'(x) = f(x)g(x)\left\{\frac{f'(x)}{f(x)} + \frac{g'(x)}{g(x)}\right\},$$

where the latter rearrangement presupposes that $f(x) \neq 0$ and $g(x) \neq 0$.
If $g(x) \neq 0$, we get the usual formula known from high school

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

It is often more convenient to compute this expression in the following way:

$$\left\{\frac{f(x)}{g(x)}\right\} = \frac{d}{dx}\left\{f(x) \cdot \frac{1}{g(x)}\right\} = \frac{f'(x)}{g(x)} - \frac{f(x)g'(x)}{g(x)^2} = \frac{f(x)}{g(x)}\left\{\frac{f'(x)}{f(x)} - \frac{g'(x)}{g(x)}\right\},$$

where the former expression often is *much easier* to use in practice than the usual formula from high school, and where the latter expression again presupposes that $f(x) \neq 0$ and $g(x) \neq 0$. Under these assumptions we see that the formulæ above can be written

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

$$\frac{\{f(x)/g(x)\}'}{f(x)/g(x)} = \frac{f'(x)}{f(x)} - \frac{g'(x)}{g(x)}.$$

Since

$$\frac{d}{dx}\ln|f(x)| = \frac{f'(x)}{f(x)}, \qquad f(x) \neq 0,$$

we also name these the *logarithmic derivatives*.

Finally, we mention the rule of **differentiation of a composite function**

$$\{f(\varphi(x))\}' = f'(\varphi(x)) \cdot \varphi'(x).$$

We first differentiate the function itself; then the insides. This rule is a 1-dimensional version of the so-called *Chain rule*.

A.4 Special derivatives.

Power like:

$$\frac{d}{dx}(x^\alpha) = \alpha \cdot x^{\alpha-1}, \qquad \text{for } x > 0, \quad \text{(extensions for some } \alpha).$$

$$\frac{d}{dx}\ln|x| = \frac{1}{x}, \qquad \text{for } x \neq 0.$$

Exponential like:

$$\frac{d}{dx}\exp x = \exp x, \qquad \text{for } x \in \mathbb{R},$$

$$\frac{d}{dx}\left(a^x\right) = \ln a \cdot a^x, \qquad \text{for } x \in \mathbb{R} \text{ og } a > 0.$$

Trigonometric:

$$\frac{d}{dx}\sin x = \cos x, \qquad \text{for } x \in \mathbb{R},$$

$$\frac{d}{dx}\cos x = -\sin x, \qquad \text{for } x \in \mathbb{R},$$

$$\frac{d}{dx}\tan x = 1 + \tan^2 x = \frac{1}{\cos^2 x}, \qquad \text{for } x \neq \frac{\pi}{2} + p\pi, p \in \mathbb{Z},$$

$$\frac{d}{dx}\cot x = -(1 + \cot^2 x) = -\frac{1}{\sin^2 x}, \qquad \text{for } x \neq p\pi, p \in \mathbb{Z}.$$

Hyperbolic:

$$\frac{d}{dx}\sinh x = \cosh x, \qquad \text{for } x \in \mathbb{R},$$

$$\frac{d}{dx}\cosh x = \sinh x, \qquad \text{for } x \in \mathbb{R},$$

$$\frac{d}{dx}\tanh x = 1 - \tanh^2 x = \frac{1}{\cosh^2 x}, \qquad \text{for } x \in \mathbb{R},$$

$$\frac{d}{dx}\coth x = 1 - \coth^2 x = -\frac{1}{\sinh^2 x}, \qquad \text{for } x \neq 0.$$

Inverse trigonometric:

$$\frac{d}{dx}\,\text{Arcsin}\,x = \frac{1}{\sqrt{1-x^2}}, \qquad \text{for } x \in \,]-1,1\,[,$$

$$\frac{d}{dx}\,\text{Arccos}\,x = -\frac{1}{\sqrt{1-x^2}}, \qquad \text{for } x \in \,]-1,1\,[,$$

$$\frac{d}{dx}\,\text{Arctan}\,x = \frac{1}{1+x^2}, \qquad \text{for } x \in \mathbb{R},$$

$$\frac{d}{dx}\,\text{Arccot}\,x = \frac{1}{1+x^2}, \qquad \text{for } x \in \mathbb{R}.$$

Inverse hyperbolic:

$$\frac{d}{dx}\,\text{Arsinh}\,x = \frac{1}{\sqrt{x^2+1}}, \qquad \text{for } x \in \mathbb{R},$$

$$\frac{d}{dx}\,\text{Arcosh}\,x = \frac{1}{\sqrt{x^2-1}}, \qquad \text{for } x \in \,]1,+\infty\,[,$$

$$\frac{d}{dx}\,\text{Artanh}\,x = \frac{1}{1-x^2}, \qquad \text{for } |x| < 1,$$

$$\frac{d}{dx}\,\text{Arcoth}\,x = \frac{1}{1-x^2}, \qquad \text{for } |x| > 1.$$

Remark A.2 The derivative of the trigonometric and the hyperbolic functions are to some extent exponential like. The derivatives of the inverse trigonometric and inverse hyperbolic functions are power like, because we include the logarithm in this class. $\Diamond$

A.5 Integration

The most obvious rules are about linearity

$$\int \{f(x) + \lambda g(x)\}\,dx = \int f(x)\,dx + \lambda \int g(x)\,dx, \qquad \text{where } \lambda \in \mathbb{R} \text{ is a constant,}$$

and about that differentiation and integration are “inverses to each other”, i.e. modulo some arbitrary constant $c \in \mathbb{R}$, which often tacitly is missing,

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

If we in the latter formula replace $f(x)$ by the product $f(x)g(x)$, we get by reading from the right to the left and then differentiating the product,

$$f(x)g(x) = \int \{f(x)g(x)\}'\,dx = \int f'(x)g(x)\,dx + \int f(x)g'(x)\,dx.$$

Hence, by a rearrangement

The rule of partial integration:

$$\int f'(x)g(x)\,dx = f(x)g(x) - \int f(x)g'(x)\,dx.$$

The differentiation is moved from one factor of the integrand to the other one by changing the sign and adding the term $f(x)g(x)$.

Remark A.3 This technique was earlier used a lot, but is almost forgotten these days. It must be revived, because MAPLE and pocket calculators apparently do not know it. It is possible to construct examples where these devices cannot give the exact solution, unless you first perform a partial integration yourself. $\diamondsuit$

Remark A.4 This method can also be used when we estimate integrals which cannot be directly calculated, because the antiderivative is not contained in e.g. the catalogue of MAPLE. The idea is by a succession of partial integrations to make the new integrand smaller. See also Chapter 4. $\diamondsuit$

Integration by substitution:

If the integrand has the special structure $f(\varphi(x))\cdot\varphi'(x)$, then one can change the variable to $y = \varphi(x)$:

$$\int f(\varphi(x))\cdot\varphi'(x)\,dx = \text{“}\int f(\varphi(x))\,d\varphi(x)\text{''} = \int_{y=\varphi(x)} f(y)\,dy.$$

Integration by a monotonous substitution:

If $\varphi(y)$ is a *monotonous* function, which maps the y-interval *one-to-one* onto the x-interval, then

$$\int f(x)\,dx = \int_{y=\varphi^{-1}(x)} f(\varphi(y))\varphi'(y)\,dy.$$

Remark A.5 This rule is usually used when we have some "ugly" term in the integrand $f(x)$. The idea is to put this ugly term equal to $y = \varphi^{-1}(x)$. When e.g. x occurs in $f(x)$ in the form $\sqrt{x}$, we put $y = \varphi^{-1}(x) = \sqrt{x}$, hence $x = \varphi(y) = y^2$ og $\varphi'(y) = 2y$. $\Diamond$

A.6 Special antiderivatives

Power like:

$$\int \frac{1}{x}\,dx = \ln|x|, \qquad \text{for } x \neq 0. \quad \text{(Do not forget the numerical value!)}$$

$$\int x^{\alpha}\,dx = \frac{1}{\alpha+1}x^{\alpha+1}, \qquad \text{for } \alpha \neq -1,$$

$$\int \frac{1}{1+x^2}\,dx = \text{Arctan}\,x, \qquad \text{for } x \in \mathbb{R},$$

$$\int \frac{1}{1-x^2}\,dx = \frac{1}{2}\ln\left|\frac{1+x}{1-x}\right|, \qquad \text{for } x \neq \pm 1,$$

$$\int \frac{1}{1-x^2}\,dx = \text{Artanh}\,x, \qquad \text{for } |x| < 1,$$

$$\int \frac{1}{1-x^2}\,dx = \text{Arcoth}\,x, \qquad \text{for } |x| > 1,$$

$$\int \frac{1}{\sqrt{1-x^2}}\,dx = \text{Arcsin}\,x, \qquad \text{for } |x| < 1,$$

$$\int \frac{1}{\sqrt{1-x^2}}\,dx = -\,\text{Arccos}\,x, \qquad \text{for } |x| < 1,$$

$$\int \frac{1}{\sqrt{x^2+1}}\,dx = \text{Arsinh}\,x, \qquad \text{for } x \in \mathbb{R},$$

$$\int \frac{1}{\sqrt{x^2+1}}\,dx = \ln(x+\sqrt{x^2+1}), \qquad \text{for } x \in \mathbb{R},$$

$$\int \frac{x}{\sqrt{x^2-1}}\,dx = \sqrt{x^2-1}, \qquad \text{for } x \in \mathbb{R},$$

$$\int \frac{1}{\sqrt{x^2-1}}\,dx = \text{Arcosh}\,x, \qquad \text{for } x > 1,$$

$$\int \frac{1}{\sqrt{x^2-1}}\,dx = \ln|x+\sqrt{x^2-1}|, \qquad \text{for } x > 1 \text{ eller } x < -1.$$

There is an error in the programs of the pocket calculators TI-92 and TI-89. The *numerical signs* are missing. It is obvious that $\sqrt{x^2-1} < |x|$ so if $x < -1$, then $x + \sqrt{x^2-1} < 0$. Since you cannot take the logarithm of a negative number, these pocket calculators will give an error message.

Exponential like:

$$\int \exp x \, dx = \exp x, \qquad \text{for } x \in \mathbb{R},$$

$$\int a^x \, dx = \frac{1}{\ln a} \cdot a^x, \qquad \text{for } x \in \mathbb{R}, \text{ og } a > 0, a \neq 1.$$

Trigonometric:

$$\int \sin x \, dx = -\cos x, \qquad \text{for } x \in \mathbb{R},$$

$$\int \cos x \, dx = \sin x, \qquad \text{for } x \in \mathbb{R},$$

$$\int \tan x \, dx = -\ln|\cos x|, \qquad \text{for } x \neq \frac{\pi}{2} + p\pi, \quad p \in \mathbb{Z},$$

$$\int \cot x \, dx = \ln|\sin x|, \qquad \text{for } x \neq p\pi, \quad p \in \mathbb{Z},$$

$$\int \frac{1}{\cos x} \, dx = \frac{1}{2} \ln\left(\frac{1+\sin x}{1-\sin x}\right), \qquad \text{for } x \neq \frac{\pi}{2} + p\pi, \quad p \in \mathbb{Z},$$

$$\int \frac{1}{\sin x} \, dx = \frac{1}{2} \ln\left(\frac{1-\cos x}{1+\cos x}\right), \qquad \text{for } x \neq p\pi, \quad p \in \mathbb{Z},$$

$$\int \frac{1}{\cos^2 x} \, dx = \tan x, \qquad \text{for } x \neq \frac{\pi}{2} + p\pi, \quad p \in \mathbb{Z},$$

$$\int \frac{1}{\sin^2 x} \, dx = -\cot x, \qquad \text{for } x \neq p\pi, \quad p \in \mathbb{Z}.$$

Hyperbolic:

$$\int \sinh x\, dx = \cosh x, \qquad \text{for } x \in \mathbb{R},$$

$$\int \cosh x\, dx = \sinh x, \qquad \text{for } x \in \mathbb{R},$$

$$\int \tanh x\, dx = \ln \cosh x, \qquad \text{for } x \in \mathbb{R},$$

$$\int \coth x\, dx = \ln |\sinh x|, \qquad \text{for } x \neq 0,$$

$$\int \frac{1}{\cosh x}\, dx = \operatorname{Arctan}(\sinh x), \qquad \text{for } x \in \mathbb{R},$$

$$\int \frac{1}{\cosh x}\, dx = 2\operatorname{Arctan}(e^x), \qquad \text{for } x \in \mathbb{R},$$

$$\int \frac{1}{\sinh x}\, dx = \frac{1}{2}\ln\left(\frac{\cosh x - 1}{\cosh x + 1}\right), \qquad \text{for } x \neq 0,$$

$$\int \frac{1}{\sinh x}\, dx = \ln\left|\frac{e^x - 1}{e^x + 1}\right|, \qquad \text{for } x \neq 0,$$

$$\int \frac{1}{\cosh^2 x}\, dx = \tanh x, \qquad \text{for } x \in \mathbb{R},$$

$$\int \frac{1}{\sinh^2 x}\, dx = -\coth x, \qquad \text{for } x \neq 0.$$

A.7 Trigonometric formulæ

The trigonometric formulæ are closely connected with circular movements. Thus $(\cos u, \sin u)$ are the coordinates of a piont P on the unit circle corresponding to the angle u, cf. figure A.1. This geometrical interpretation is used from time to time.

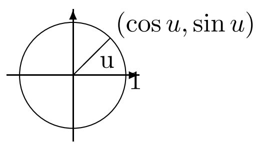

Figure 42: The unit circle and the trigonometric functions.

The fundamental trigonometric relation:

$$\cos^2 u + \sin^2 u = 1, \qquad \text{for } u \in \mathbb{R}.$$

Using the previous geometric interpretation this means according to *Pythagoras's theorem*, that the point P with the coordinates $(\cos u, \sin u)$ always has distance 1 from the origo $(0,0)$, i.e. it is lying on the boundary of the circle of centre $(0,0)$ and radius $\sqrt{1} = 1$.

Connection to the complex exponential function:

The *complex exponential* is for imaginary arguments defined by

$$\exp(\mathrm{i}\,u) := \cos u + \mathrm{i}\ \sin u.$$

It can be checked that the usual functional equation for exp is still valid for complex arguments. In other word: The definition above is extremely conveniently chosen.

By using the definition for $\exp(\mathrm{i}\,u)$ and $\exp(-\mathrm{i}\,u)$ it is easily seen that

$$\cos u = \frac{1}{2}(\exp(\mathrm{i}\,u) + \exp(-\mathrm{i}\,u)),$$

$$\sin u = \frac{1}{2i}(\exp(\mathrm{i}\,u) - \exp(-\mathrm{i}\,u)),$$

.

Moivre's formula: By expressing $\exp(\mathrm{i}nu)$ in two different ways we get:

$$\exp(\mathrm{i}nu) = \cos nu + \mathrm{i}\sin nu = (\cos u + \mathrm{i}\sin u)^n.$$

Example A.1 If we e.g. put $n = 3$ into Moivre's formula, we obtain the following typical appliction,

$$\begin{aligned}\cos(3u) + \mathrm{i}\sin(3u) &= (\cos u + \mathrm{i}\sin u)^3\\ &= \cos^3 u + 3\mathrm{i}\cos^2 u \cdot \sin u + 3\mathrm{i}^2 \cos u \cdot \sin^2 u + \mathrm{i}^3 \sin^3 u\\ &= \{\cos^3 u - 3\cos u \cdot \sin^2 u\} + \mathrm{i}\{3\cos^2 u \cdot \sin u - \sin^3 u\}\\ &= \{4\cos^3 u - 3\cos u\} + \mathrm{i}\{3\sin u - 4\sin^3 u\}\end{aligned}$$

When this is split into the real- and imaginary parts we obtain

$$\cos 3u = 4\cos^3 u - 3\cos u, \qquad \sin 3u = 3\sin u - 4\sin^3 u. \quad \Diamond$$

Addition formulæ:

$$\sin(u+v) = \sin u \cos v + \cos u \sin v,$$
$$\sin(u-v) = \sin u \cos v - \cos u \sin v,$$
$$\cos(u+v) = \cos u \cos v - \sin u \sin v,$$
$$\cos(u-v) = \cos u \cos v + \sin u \sin v.$$

Products of trigonometric functions to a sum:

$$\sin u \cos v = \frac{1}{2}\sin(u+v) + \frac{1}{2}\sin(u-v),$$
$$\cos u \sin v = \frac{1}{2}\sin(u+v) - \frac{1}{2}\sin(u-v),$$
$$\sin u \sin v = \frac{1}{2}\cos(u-v) - \frac{1}{2}\cos(u+v),$$
$$\cos u \cos v = \frac{1}{2}\cos(u-v) + \frac{1}{2}\cos(u+v).$$

Sums of trigonometric functions to a product:

$$\sin u + \sin v = 2\sin\left(\frac{u+v}{2}\right)\cos\left(\frac{u-v}{2}\right),$$
$$\sin u - \sin v = 2\cos\left(\frac{u+v}{2}\right)\sin\left(\frac{u-v}{2}\right),$$
$$\cos u + \cos v = 2\cos\left(\frac{u+v}{2}\right)\cos\left(\frac{u-v}{2}\right),$$
$$\cos u - \cos v = -2\sin\left(\frac{u+v}{2}\right)\sin\left(\frac{u-v}{2}\right).$$

Formulæ of halving and doubling the angle:

$$\sin 2u = 2\sin u \cos u,$$
$$\cos 2u = \cos^2 u - \sin^2 u = 2\cos^2 u - 1 = 1 - 2\sin^2 u,$$
$$\sin\frac{u}{2} = \pm\sqrt{\frac{1-\cos u}{2}} \qquad \text{followed by a discussion of the sign,}$$
$$\cos\frac{u}{2} = \pm\sqrt{\frac{1+\cos u}{2}} \qquad \text{followed by a discussion of the sign,}$$

A.8 Hyperbolic formulæ

These are very much like the trigonometric formulæ, and if one knows a little of Complex Function Theory it is realized that they are actually identical. The structure of this section is therefore the same as for the trigonometric formulæ. The reader should compare the two sections concerning similarities and differences.

The fundamental relation:

$$\cosh^2 x - \sinh^2 x = 1.$$

Definitions:

$$\cosh x = \frac{1}{2}\left(\exp(x) + \exp(-x)\right), \qquad \sinh x = \frac{1}{2}\left(\exp(x) - \exp(-x)\right).$$

"Moivre's formula":

$$\exp(x) = \cosh x + \sinh x.$$

This is trivial and only rarely used. It has been included to show the analogy.

Addition formulæ:

$$\sinh(x+y) = \sinh(x)\cosh(y) + \cosh(x)\sinh(y),$$

$$\sinh(x-y) = \sinh(x)\cosh(y) - \cosh(x)\sinh(y),$$

$$\cosh(x+y) = \cosh(x)\cosh(y) + \sinh(x)\sinh(y),$$

$$\cosh(x-y) = \cosh(x)\cosh(y) - \sinh(x)\sinh(y).$$

Formulæ of halving and doubling the argument:

$$\sinh(2x) = 2\sinh(x)\cosh(x),$$

$$\cosh(2x) = \cosh^2(x) + \sinh^2(x) = 2\cosh^2(x) - 1 = 2\sinh^2(x) + 1,$$

$$\sinh\left(\frac{x}{2}\right) = \pm\sqrt{\frac{\cosh(x)-1}{2}} \qquad \text{followed by a discussion of the sign,}$$

$$\cosh\left(\frac{x}{2}\right) = \sqrt{\frac{\cosh(x)+1}{2}}.$$

Inverse hyperbolic functions:

$$\operatorname{Arsinh}(x) = \ln\left(x + \sqrt{x^2+1}\right), \qquad x \in \mathbb{R},$$

$$\operatorname{Arcosh}(x) = \ln\left(x + \sqrt{x^2-1}\right), \qquad x \geq 1,$$

$$\operatorname{Artanh}(x) = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right), \qquad |x| < 1,$$

$$\operatorname{Arcoth}(x) = \frac{1}{2}\ln\left(\frac{x+1}{x-1}\right), \qquad |x| > 1.$$

A.9 Complex transformation formulæ

$$\cos(\mathrm{i}x) = \cosh(x), \qquad \cosh(\mathrm{i}x) = \cos(x),$$

$$\sin(\mathrm{i}x) = \mathrm{i}\,\sinh(x), \qquad \sinh(\mathrm{i}x) = \mathrm{i}\,\sin x.$$

A.10 Taylor expansions

The generalized binomial coefficients are defined by

$$\binom{\alpha}{n} := \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{1\cdot 2\cdots n},$$

with n factors in the numerator and the denominator, supplied with

$$\binom{\alpha}{0} := 1.$$

The Taylor expansions for *standard functions* are divided into *power like* (the radius of convergency is finite, i.e. $= 1$ for the standard series) and*exponential like* (the radius of convergency is infinite).

Power like:

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n, \qquad |x| < 1,$$

$$\frac{1}{1+x} = \sum_{n=0}^{\infty} (-1)^n x^n, \qquad |x| < 1,$$

$$(1+x)^n = \sum_{j=0}^{n} \binom{n}{j} x^j, \qquad n \in \mathbb{N}, x \in \mathbb{R},$$

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n, \qquad \alpha \in \mathbb{R} \setminus \mathbb{N}, |x| < 1,$$

$$\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n}, \qquad |x| < 1,$$

$$\operatorname{Arctan}(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}, \qquad |x| < 1.$$

Exponential like:

$$\exp(x) = \sum_{n=0}^{\infty} \frac{1}{n!} x^n, \qquad x \in \mathbb{R}$$

$$\exp(-x) = \sum_{n=0}^{\infty} (-1)^n \frac{1}{n!} x^n, \qquad x \in \mathbb{R}$$

$$\sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n+1)!} x^{2n+1}, \qquad x \in \mathbb{R},$$

$$\sinh(x) = \sum_{n=0}^{\infty} \frac{1}{(2n+1)!} x^{2n+1}, \qquad x \in \mathbb{R},$$

$$\cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n)!} x^{2n}, \qquad x \in \mathbb{R},$$

$$\cosh(x) = \sum_{n=0}^{\infty} \frac{1}{(2n)!} x^{2n}, \qquad x \in \mathbb{R}.$$

A.11 Magnitudes of functions

We often have to compare functions for $x \to 0+$, or for $x \to \infty$. The simplest type of functions are therefore arranged in an hierarchy:

1) logarithms,

2) power functions,

3) exponential functions,

4) faculty functions.

When $x \to \infty$, a function from a higher class will always dominate a function form a lower class. More precisely:

A) A *power function* dominates a *logarithm* for $x \to \infty$:

$$\frac{(\ln x)^\beta}{x^\alpha} \to 0 \qquad \text{for } x \to \infty, \quad \alpha, \beta > 0.$$

B) An *exponential* dominates a *power function* for $x \to \infty$:

$$\frac{x^\alpha}{a^x} \to 0 \qquad \text{for } x \to \infty, \quad \alpha, a > 1.$$

C) The *faculty function* dominates an *exponential* for $n \to \infty$:

$$\frac{a^n}{n!} \to 0, \qquad n \to \infty, \quad n \in \mathbb{N}, \quad a > 0.$$

D) When $x \to 0+$ we also have that a *power function* dominates the *logarithm:*

$$x^\alpha \ln x \to 0-, \qquad \text{for } x \to 0+, \quad \alpha > 0.$$