

Contents

Introduction.....	3
1: Some basic mathematics	4
2: Why nature loves logarithms	9
3: Proportionalities.....	15
4: Basic functions and their application in biology.....	23
5. Handling a changing world.....	32
6: Summing up.....	41
6.1: The Newton approximation	48
7: The sum of infinities	49
Literature.....	58
Online archives and textbooks	59
Mathematical software.....	60

Introduction

“The book of nature is written in the language of mathematics”, Galileo wrote 400 years ago. Still 15 years ago the mathematical skills of most biology students ended with school mathematics enriched with a little basic statistics. The third industrial revolution with its powerful mathematics and statistics packages changed this situation dramatically. The mathematization of our world did not leave out biology departments and the rapid transformation of biology from a merely descriptive to an explanatory science set new standards with regard to data analysis and modelling skills. Today, a sound mathematical education (together with the knowledge of at least one computer language) becomes more and more a necessity to find a job after study.

The following text is the first part of a lecture in basic mathematics and statistics for biologists. The lecture contains what might be considered an international standard of basic knowledge although many readers will surely miss important branches. However, a one year course that has to deal with mathematics modelling must be to a certain extent eclectic. Emphasis was especially paid to basic mathematical techniques and principles of biological modelling. Many examples are included that show how to program simple tasks with a spreadsheet program and how to use advanced mathematics software. The text does not repeat school mathematics. Therefore, basic algebra and especially geometry are as well missing as approximation techniques and integral solving. Today, math programs do such jobs for us and the aim of this introductory course is more to teach how to interpret their results.

The following text is not a textbook. It is intended as a script to present the contents of the lecture in a condensed form. There is no need to write a textbook again. Today, the internet took over many former tasks textbooks had. The end of this text contains therefore a small overview over important internet pages where students can find mathematics glossaries, textbooks, and program collections.

1. Some basic mathematics

All mathematics begins with counting. Then **natural numbers** 1, 2, 3, ... have an intuitive appeal and every child begins calculating with counting and simple adding. Natural numbers appear natural to us because of the intuitive way of learning them. Even some animals like parrots, ravens, or chimps are able to count or even to do some simple calculations.

Mathematicians denote the natural numbers with the symbol \mathbb{N} . \mathbb{N} is defined in the range of 0 to infinity. Everybody knows of course a first extension of \mathbb{N} . It is the set of **whole numbers** \mathbb{Z} (... , -3, -2, -1, 0, 1, 2, 3, 4, ...), defined in the range from $-\infty$ to $+\infty$. Hence, mathematicians write

$$\mathbb{N} \subset \mathbb{Z}$$

that means the **set of the natural numbers** is contained in the set of the whole numbers. Note that \mathbb{N} is defined by the most elementary operation, by adding. The operation of subtracting is also possible within \mathbb{N} but with an important limitation, $5 - 3$ is defined but not $3 - 5$. If we want to overcome this limitation we need \mathbb{Z} . In the same way we can multiply within \mathbb{Z} but for the next operation, division, we have again a limitation because $3 / 4$ is not defined within \mathbb{Z} . Hence we need the next extension, the set of the **rational numbers** and note

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}$$

Again, another operation is fully defined within \mathbb{Q} , the power. You can compute $3.3^{5.1}$ within \mathbb{Q} . But taking the root is not fully defined. From school we know that $\sqrt{2}$ is a so-called **irrational number** because it cannot be written in the form a / b where a and b are both elements of \mathbb{Z} . To define operations like $\sqrt{2}$ we need again a new class of numbers, the **real numbers**

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$$

A	α	Alpha
B	β	Beta
Γ	γ	Gamma
Δ	δ	Delta
E	ε	Epsilon
Z	ζ	Zeta
H	η	Eta
Θ	θ	Theta
I	ι	Jota
K	κ	Kappa
Λ	λ	Lambda
M	μ	My
N	ν	Ny
Ξ	ξ	Xi
O	\omicron	Omikron
Π	π	Pi
P	ρ	Rho
Σ	σ	Sigma
T	τ	Tau
Y	υ	Ypsilon
Φ	ϕ	Phi
X	χ	Chi
Ψ	ψ	Psi
Ω	ω	Omega

Hence 5 is an element of \mathbb{N} , of \mathbb{Z} , of \mathbb{Q} , and of \mathbb{R} .

$$5 \in \mathbb{R}$$

Of course, all of this should be known from school. It shows us two important things. First, mathematicians use a symbolic language to denote their objects. \mathbb{N} , \mathbb{Z} , \mathbb{Q} , and \mathbb{R} are symbols. Second, we started with an intuitive approach by defining objects via basic mathematical operations, adding, subtracting, taking the power and roots.

The number of symbols is limited. Therefore, mathematicians have to use different sign sets and alphabets to handle with their objects. Most often used is the greek alphabet given in Box 1. Box 2

Set theory

\in is element of
 \notin is not element of
 \subset is part of

Logic

\wedge and
 \vee or
 \neg not
 \rightarrow if ... then
 \leftrightarrow exactly if... then

Algebra

$=$ equals
 \approx approximately equals
 $<$ less than
 $>$ larger than
 \propto, \sim proportional to
 $|a|$ absolute value of a

shows some often used operators.

The intuitive way is not the only approach to start. Another approach is to start with a set of suppositions, so-called **axioms**. These are sentences of a **formal language** that contain basic relations between mathematical objects. These relations implicitly define these objects. The first mathematician who based a part of the mathematical science, the geometry, on a set of axioms, was the Greek mathematician Euclid (about 365?-300? Ad). His textbook *Elements* was used in schools until the 19th century. The German mathematician David Hilbert (1862-1943) intended to base all parts of mathematics on axiomatic systems, a project that had an immense influence on today's way of making science and that is called **metamathematic**.

Our natural numbers can also be introduced via axioms. The system of the natural numbers of the Italian mathematician Giuseppe Peano (1858 - 1932) and the German Richard Dedekind (1831 - 1916) looks as follows

1. *Zero is a natural number*
2. *Every natural number n has a follower n' defined as $n+1$.*
3. *Zero does not follow any natural number*
4. *If n' follows n and m' follows n and $m = n$ then $n' = m'$.*
5. *Every set that contains zero, and with a number n also has its follower n' contains the whole set of natural numbers (the principle of exhaustive induction).*

We see that axioms need not to be intuitively clear or even self evident. On first look the Peano Dedekind axioms appear quite odd. However, these five axioms allow to derive basic rules of calculations which are defined within the natural numbers. What is for instance $2 + 2$? Axioms 2 and 4 define a series of elements related by a strict larger or smaller relation. These elements are termed by us 0, 1, 2, 3, etc. Hence numbers are in our axiomatic system only the names for those elements defined by the axioms. Axiom 2 tells that 2 is 1' hence 0". Therefore, $2 + 2$ is 2" of 0", the second follower of the second follower of zero. This is 0"". We name the fourth follower of zero four.

This looks very abstract but shows us important features every scientific axiomatic system must have. First of all axioms must be **consistent** (they must not lead to contradictions). All axioms must be **independent** of each other. Axioms must implicitly define basic objects and relations. For instance, the above axioms of the natural numbers define as a whole the natural numbers, but they do not define the exact meaning of follower. However, no axiomatic system is complete, it is not that it contains all axioms being necessary to derive all mathematical theorems that can be formulated with the language used to formulate the axioms. The metamathematical investigations during the last century, especially the famous theorem of Kurt Gödel (German mathematician, 1906-1978), showed that every axiomatic system can be extended by new axioms, that are independent of the former, to define a new set of objects of which the former objects are part.

Beside axioms, mathematicians need something else, rules that make it possible to derive mathematical sentences, **theorems**, from others. These rules allow us to prove theorems. A **proof** is therefore in our context nothing else than deriving a theorem from the set of axioms using this set of rules. **Hence, axioms themselves cannot be proven.**

Mathematics knows four such rules of derivation (maybe there are even more)

1. The **modus ponens**: If a theorem A is true and from A follows B then B is also true.
 $A; A \rightarrow B; B$.

2. The **modus tollens**: If B follows from A and B is not true, then A is also not true.
 $A \rightarrow B; \neg B; \neg A.$
3. **Tertium non datur**: Either A is true or A is false, a third possibility is impossible.
 $A \vee \neg A$
4. The **principle of complete induction**: Assume a theorem is true for a natural number n. If it is true for any number m larger than n and with necessity also for the follower m' of m than it is true for all natural numbers equal or larger than n.
 $A(k), A(n \geq k) \rightarrow A(n+1); A(i \geq k)$

It is not necessary to apply all four rules. In particular the third rule, the tertium non datur, has provoked much discussion, and it is still an open question whether this rule is at all necessary (although its acceptance makes things much easier). A series of theories in logic and natural sciences (quantum theory for instance) explicitly negate the tertium non datur leaving the majority of elementary theorems still valid.

Mathematical objects are connected via operators, which are implicitly defined by axioms. Such connections are called **terms**. $5 + 7$ is a term. $\sqrt{2}$ is also a term. Through comparing operators (+, <, or >) we can combine terms to form equations or comparisons.

In essence, mathematics is a formal science that deals with objects that can be combined by operations. By this structures are formed. **Mathematics is the science of structures**. In this interpretation these structures need not to be real. Even more, nothing is said about any relation to reality. How can we then describe reality with the aim of mathematics? Because we always simplify. Our senses and our brain see the world outside and inside us according to inherent rules (emerged during evolution as a representation of external structures through internal ones). These **a priori** rules identify therefore external structures that can be described by mathematics. In doing science (for instance biology) we seek such structures. And we have to describe them. In most cases we will do this in a technical language. By this we build a **model**.

We construct models in order to deal with complex processes. Models allow us to understand these processes and to infer the variables, the so called **drivers**, that influence them. Models also allow us to predict future states. This predicting power is important for instance in models of climate change, ecosystem functioning, population growth, or enzymatic activity and genetics, fields where appropriate direct observations or experiments are often impossible. Biology changed from a mainly descriptive science to an explanatory science. **To explain things means to be able to model them**.

A very simple example. Only a few insects have reproductive cycles longer than one year. In a series of cicadas however imagines are produced every 7, 13, or even 17 years (in the American cikade *Magicicada septendecim*).



Magicicada septendecim
 Photo by USA National Arboretum

How can we explain such odd numbers? It was found that these species live in very predictable habitats with rather constant amounts of resources. Their main predators are vertebrates, birds or reptiles with reproductive cycles of two, four or six years. Now we make a very simple (and of course simplified) model. We set predator abundances in reproductive years to values being two times as high as in non-reproductive years (Fig. 1.1). The sum of all predator abundances is then the predation pressure on our cicadas. This can be modelled by a

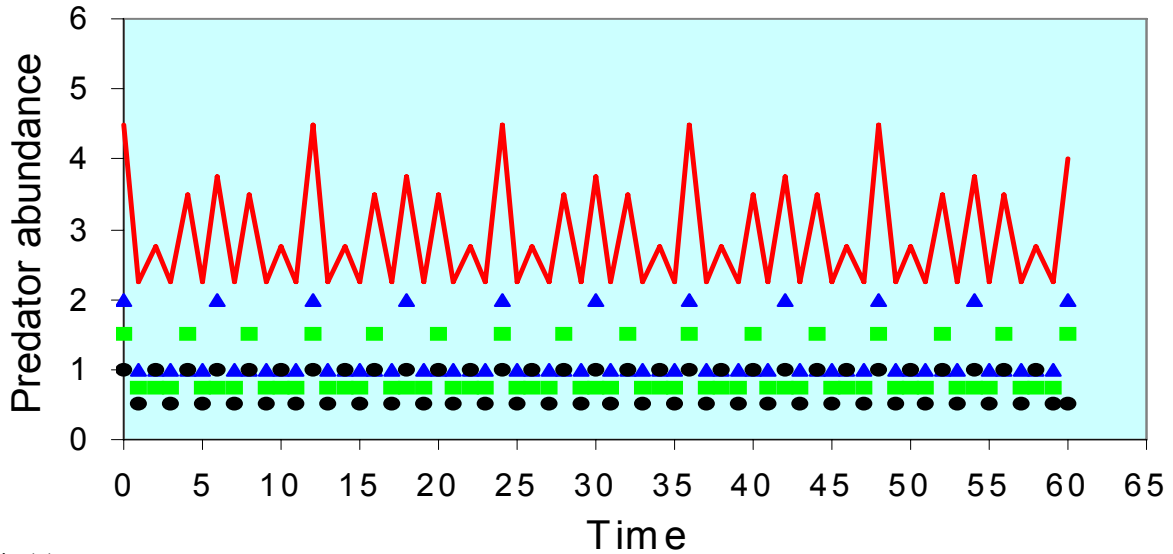


Fig. 1.1

simple Excel spreadsheet. Now look at the Figure 1.1. Low total predator abundances occur every 3, 5, 7, 9, 11, 13, 17, 19, etc. years.

Of course, predators do not have stable abundances and cicdads have more than two predators. We have to improve our model. For instance, we could introduce more predators or even variable numbers of

	A	B	C	D	E
1	Generation	Predator A	Predator B	Predator C	Sum of predator densities
2	0	1	1.5	2	4.5
3	1	0.5	0.75	1	2.25
4	2	1	0.75	1	2.75
5	3	0.5	1.5	1	3
6	4	1	0.75	2	3.75
7	5	0.5	0.75	1	2.25
8	+A7+1	+B6	+C5	+D4	+SUMA(B8:D8)
1	A	B	C	D	E
1	Generation	Predator A	Predator B	Predator C	Sum
2	1	=2*LOS()	=3*LOS()	=4*LOS()	=SUMA(M34:O34)
3	=A2+1	=B2*LOS()	=C2*LOS()	=D2*LOS()	=SUMA(M35:O35)
4	=A3+1	=2*LOS()	=C2*LOS()	=D2*LOS()	=SUMA(M36:O36)
5	=A4+1	=B2*LOS()	=3*LOS()	=D2*LOS()	=SUMA(M37:O37)
6	=A5+1	=2*LOS()	=C2*LOS()	=4*LOS()	=SUMA(M38:O38)

predators assigned by the model. We can also assign variable abundances. By this we make our model more realistic but also more complicated. Fig. 1.2 shows an example of such a simulation. Three predator species with 1, 2, and 4 year cycles were assigned abundances generated by the Excel build in random number generator. We see low predator abundances at prime numbers like 7, 13, 17, or 19. These are exactly the cycles found in nature and we simulated (modelled) this pattern.

Above we used a simulation and a graphical representation of our model. More often it is necessary to state our models explicitly. We have to use sets of mathematical equations. For instance, a very simple model

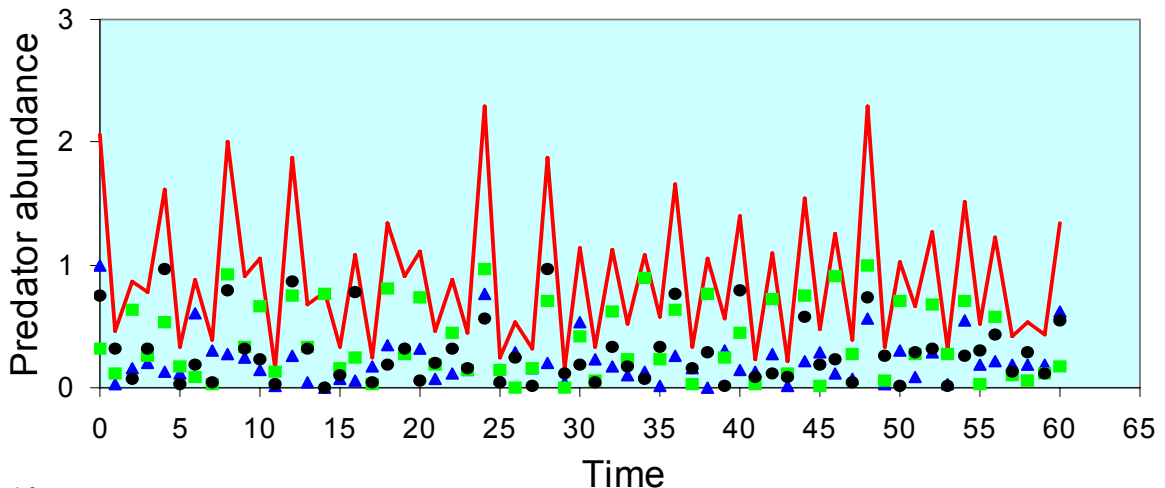
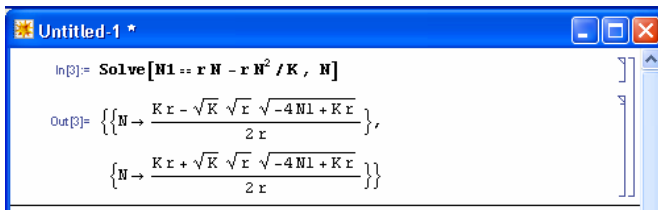


Fig. 1.2

describes the increase of bacterial populations in a constant environment.



$$\Delta N = rN - r \frac{N^2}{K}$$

where ΔN is the increase, N the actual population size and r the species specific rate of increase (fecundity). K denotes the upper limit in

population size caused by resource limitation. We are interested in the population size at a given (measured) rate of increase ΔN . Hence, we have to solve this equation for N . This should be known from school although providing the general solution might be quite time consuming for mathematically unskilled people. Fortunately, such technical things are today done by specialised mathematics programs. These are able to provide not only numerical solutions but they can also handle symbolic expressions and solve various types of equations, compute derivatives and integrals.

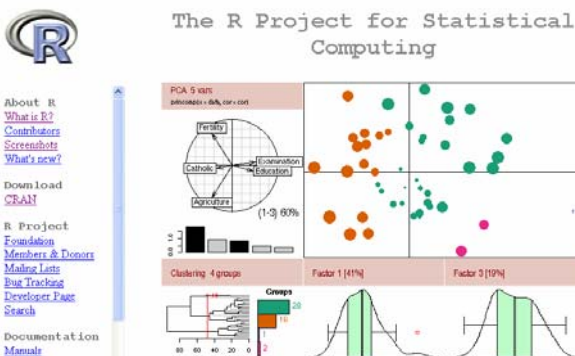


compute derivatives and integrals.

They are also able to simplify complex terms. Three important programs are *Mathematica*, *Maple*, and *Matlab*. In



statistics *Statistica*, *SAS*, *Systat*, or *SPSS* are very popular. Recently the *R-project* also gained popularity particularly among biologists. In this lecture we will mainly use the programs *Mathematica*



and the *Excel* add ins *Matrix* and *PopTools*. Above a *Mathematica* solution of our quadratic function is shown.

In our model we still would be able to solve the equation by hand. However, most biological models are more complex and in these cases math programs like *Maple* or *Mathematica* are indispensable even for trained mathematicians. Fortunately, they are now avail-

able in all university departments and larger corporations. Hence, the knowledge how to apply them is today an integrate part of any mathematics lecture. Additionally, mathematical education has also to include knowledge how to develop at least simple models using classical spreadsheet programs like *Excel*.

2. Why nature loves logarithms

Assume a snail is 10 m apart from its feeding plant and creeps with a constant speed of 0.5 m / hour. Of

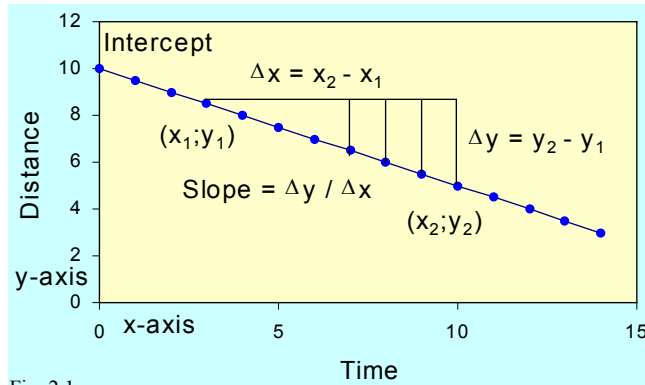


Fig. 2.1

course, after 1 hour it will be only 9.5 m apart, after 4 hours 8 m and so on. When does the snail reach the plant? Of course, after 20 hours. Surely, this is a very silly example but it provides us with very important possibilities how to describe and to visualize such a process.

Look at the Figure beside (Fig. 2.1). We use a so-called **Cartesian coordinate system** with two rectangular axes, a y- and a x-axis (after the French philosopher and mathematician Rene Des-

cartes; 1596-1650). Now we plot the distance on the y-axis and the time on the x axis (a scatter plot) . The result is linear relationship between distance and time. Such a simple linear relationship is called a **proportional relationship**, in this special case an inverse proportional relationship. Distance is inverse proportional to time. At time 0 we have the initial distance, this point is called the **intercept**. However, for a mathematical description of the whole process we need another value, the **slope**. The slope describes how fast the snail will reach the plant. The steeper the slope is, the faster the snail will reach the plant. A convenient definition of the slope is to use the quotient $\Delta y / \Delta x$ of two data points (denoted as $(x_1; y_1)$ and $(x_2; y_2)$) because simple geometry tells us that in this case of proportionality this quotient will be constant (why?). So, we can describe the whole process as follows: At time 0 the distance is 10 m. At time 1 the distance is 10 m—0.5*1 m, because the snail creeps 0.5 m / hour. At time 2 the distance is 10 m—0.5*2 m, because the snail creeps 1 m in 2 hours. And so on. In general, we have distance = initial distance – speed * time or in a more mathematical language

$$y = \text{intercept} + \text{slope} * x$$

$$y = a + mx$$

(2.1)

If m is negative (as in our case) the proportionality is inverse or indirect, for a positive slope y and x are directly proportional.

Our slope definition gives us immediately another relationship

$$\text{Slope} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{y_2 - y}{x_2 - x}$$

(2.2)

If we now want to infer the intercept we have to set x to 0. y gives then the intercept. We get

$$\text{Intercept} = y = y_2 - \frac{y_2 - y_1}{x_2 - x_1} * x_2 = y_2 - mx_2$$

(2.3)

The latter equation is a convenient way to infer slope and intercept of a linear function from two data points.

Of course, our example is a very simple one. But it acquainted us with a very natural and intuitive way to describe the relation between two variables x and y, the time and the distance. Unfortunately, such simple pro-

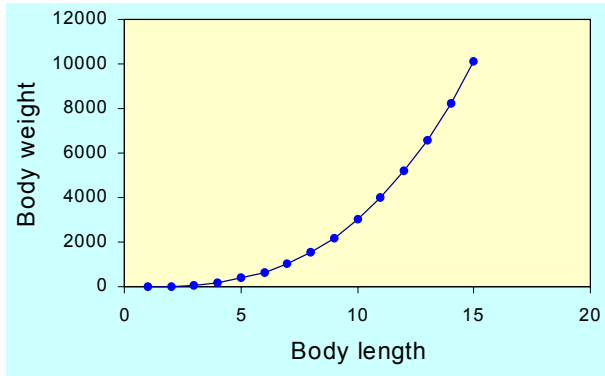
portional relationships are quite uncommon in nature. Much more common are (seemingly) more complicated relationships like the following. The volume of a cube equals the cube's length taken to the third power: $V \propto L^3$ (read V is proportional to the third power of L). The volume of an area scales to its diameter by $A \propto L^2$. For most animals and plants a similar relationship holds but the exponents are lower. For instance, insect body weights scale to body length with $W \propto L^{2.6}$. In other words, the body weight is only roughly proportional to the third power of body length; mathematically speaking

$$V \propto L^{2.6} \tag{2.4}$$

Such a relationship is shown in the Figure below (Fig. 2.2). Now we can't define an intercept or a slope of this function. The slope continuously changes and setting L to 0 is impossible. What to do?

We have to transform our equation in order to linearize it. Then we can treat it like our linear function before. In Excel axes can be rescaled logarithmically as shown below.

For this task we have to remember what a logarithm is. A logarithm is that number with which we have



to take another number (the base) to the power to get a third number. Hence, the logarithm of body weight V to base 10 is that number that fulfils the equation $V = 10^x$. In general

$$y = a^x \rightarrow x = \log_a y$$

$$x = a^{\log_a(x)} \tag{2.5}$$

Now, we have to remember some simple rules for calculations with exponents:

$$a^{x+y} = a^x * a^y; (ab)^x = a^x b^x; a^0 = 1$$

$$a^{-x} = \frac{1}{a^x}; a^{(xy)} = (a^x)^y; \sqrt[x]{a} = a^{\frac{1}{x}}$$

but

$$a^{(x^y)} \neq (a^x)^y$$

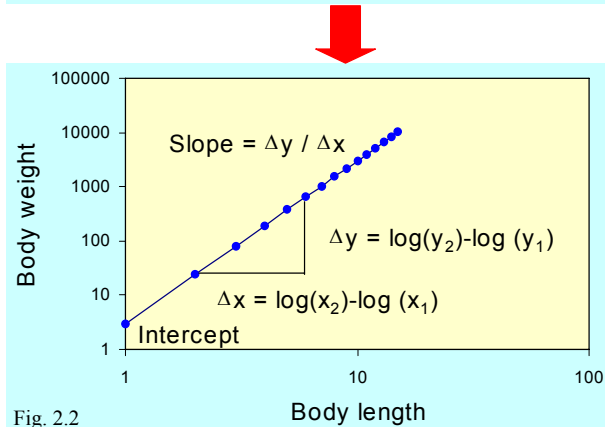


Fig. 2.2

From our definitions we get some simple computation rules

$$x * y = a^{\log_a(x)} * a^{\log_a(y)} = a^{\log_a(x) + \log_a(y)}$$

therefore

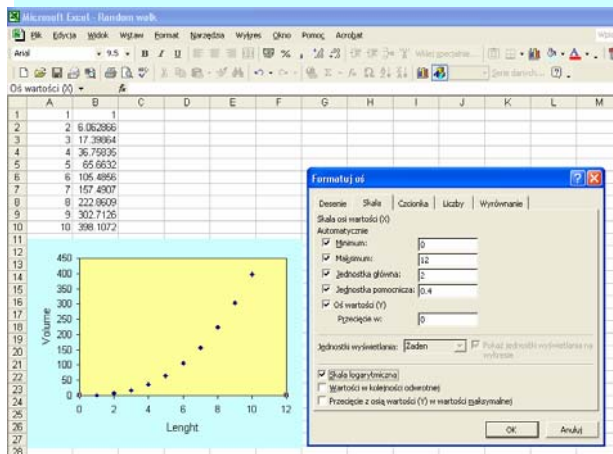
$$\log(x * y) = \log(x) + \log(y) \tag{2.6}$$

With the same logic we get

$$\log(x / y) = \log(x) - \log(y)$$

and

$$\log(x^y) = y \log(x) \tag{2.7}$$



We also notice two special cases

$$\log(1) = 0$$

and

$$\log_a(a) = 1$$

Most often used are logarithms to base 10 (called decimal logarithms), to base 2 (called binary logarithms, $\log_2(x) = \text{lb}(x)$), and to base e (called **natural or Neper logarithms**, $\log_e(x) = \ln(x)$). In particular the latter are of major importance in the natural sciences and have become a standard in the scientific literature. What is e? e is a curious number that once Neper defined from the following sum

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} \dots = \sum_{i=0}^{\infty} \frac{x^i}{i!} \tag{2.8}$$

For x = 1 we get

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} \dots = \sum_{i=0}^{\infty} \frac{1}{i!} \tag{2.9}$$

N! (read n factorial) is for natural numbers defined as the product of all numbers i from i = 1 to n. The Greek sign Σ (Sigma) is the mathematical shortage for a sum from i = 1 to n.

$$n! = \prod_{i=1}^n i = e^{\sum_{i=1}^n \ln(i)} \tag{2.10}$$

So, $5! = 1*2*3*4*5=120$; $0! = 1$ and $1! = 1$. The Greek sign Π (Pi) is the mathematical shortage for a product from i = 1 to n.

The factorial leads very fast to high numbers, for instance $10! = 3628800$, but $20!$ is already 2432902008176640000. For many applications and for analytical computation it is often convenient to use an approximation to the factorial. Such an approximation gives Stirling's equation that states

$$n! = n^n e^{-n} \sqrt{2n\pi} e^{\frac{\varepsilon(n)}{12n}} \tag{2.11}$$

In this equation ε is a number between 0 and 1. Note that $\varepsilon(n)$ (read epsilon of n) defines a function that depends on n and is not a product. For larger n the last term $e^{\varepsilon(n)/12n}$ is nearly 1 and we may approximate

$$n! \approx n^n e^{-n} \sqrt{2n\pi}$$

Stirling's approximation leads very fast to good approximations. For instance $10! = 3628800$

The approximation gives 3598693. The error is therefore only 0.8%.

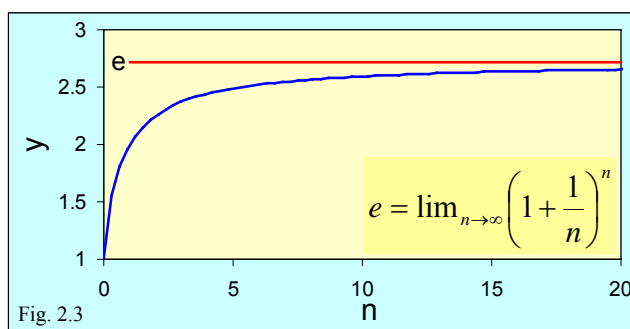


Fig. 2.3

Later we will see that the above definition of e is equivalent to another form to construct e:

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n \tag{2.12}$$

This definition uses a so-called **limes**, a boundary to which a function or series goes as some variable goes to some defined value or - as in this case - to

infinity. Figure 2.3 shows a plot of equation 2.12. We see that the curve slowly and **asymptotically** approaches but never reaches e.

e is an irrational number, a number like π that cannot be given by a rationale of the form a / b with a and b being rationale numbers. From eq. 2.12 we easily compute e

$$e \approx 2.71828183\dots$$

Often (especially in computer programs) e^x is written as $\exp(x)$. You must not confuse this with the program notation $a \text{ e}00n$ (exponent n) that is equivalent to $a10^n$.

Sometimes, we have to transform one type of logarithm into another type. This can easily be done from the following scheme:

$$x = a^{\log_a(x)} = b^{\log_b(a) * \log_a(x)} = b^{\log_b(x)}$$

hence

$$\log_b(x) = \log_b(a) * \log_a(x)$$

(2.13)

Example: $\ln(100)$ is approximately 4.605. $\log_{10}(100)$ is therefore $\log_{10}(e) * 4.605 = 0.434 * 4.605 = 1.999 \approx 2$.

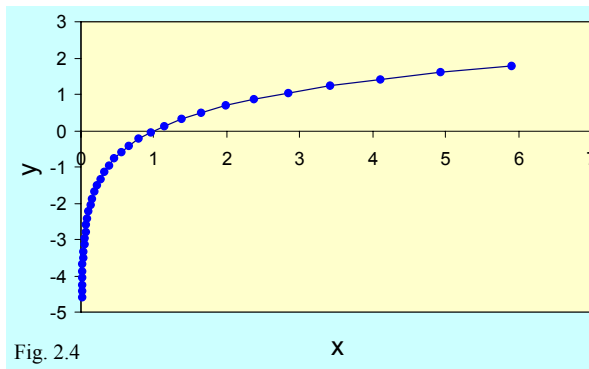


Fig. 2.4

The next Figure 2.4 shows a plot of $y = \ln(x)$ against x. We see that if x reaches 0 y goes to $-\infty$. Mathematically speaking

$$\lim_{x \rightarrow 0} \ln(x) = -\infty$$

For $x = 1$ y is 0. This is called the **root of the function**. The root of $y = \ln(x)$ is therefore $x = 1$. The function has no upper boundary. For $x \rightarrow \infty$ goes $y = \ln(x) \rightarrow \infty$.

$\rightarrow \infty$.

Above we saw that rescaling a variable results in a power function relationship (Eq. 2.4). There are many other ways to generate such non-linear relationships.

Look at the next Figure (Fig. 2.5). The numbers of individuals of species (their **abundances**) are most often unforeseeable because of the many independent factors acting on the populations. The left part of the Figure shows such abundances per unit of area (the **densities**) of 100 species, which were adjusted to a range of 0 to 1. We are interested in the fraction of species that had very low densities to estimate how many species are at the risk of going extinct. Let this fraction be p. In the next generation, these species have again random density fluctuations independent of the foregoing generations. Again, a fraction of p species will have low densities. In the next generations this process repeats. Therefore, in the second generation the fraction of species expected to have again low densities is $p * p$, in the third generation this fraction will be $p * p * p$ and so on. The fraction of

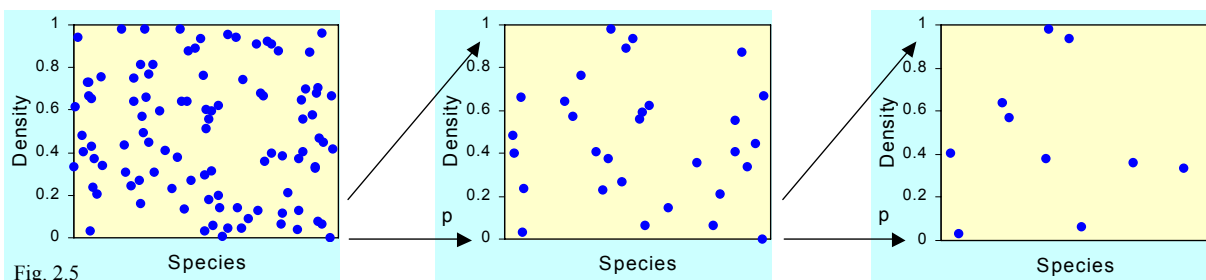


Fig. 2.5

species having low abundances remains therefore constant, but we can't foresee which species will have low densities. The whole process does not depend on any previous state. Such processes are very common in biology. For instance mutation events, radioactive decay, events of species extinctions, or hormone concentrations can be described by them.

In general, the proportion of species having constantly low densities is

$$S_i = S_0 p^i$$

where S_0 denotes the initial number of species. Using again natural logarithms and defining a value $k = -\ln(p)$ gives

$$S_i = S_0 e^{-ki} \tag{2.14}$$

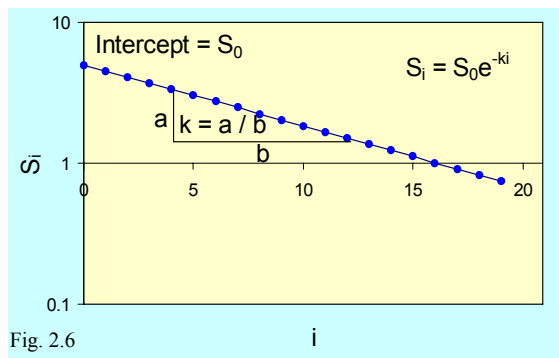


Fig. 2.6

We can linearize the equation by taking the logarithms and get

$$\ln(S_i) = \ln(S_0) - k * i$$

The next Figure (Fig. 2.6) shows a plot of S_i against i using a semi-logarithmic coordinate system. The y-axis has a logarithmic scale. Now the exponential function appears to be a straight line and the intercept at the y-axis gives us the value of S_0 . The slope of the line equals the factor k .

Of major importance is the value of i when S_i is exactly $S_0/2$, the so-called **half-time value**. We get

$$\frac{0.5 S_0}{S_0} = \frac{1}{2} = e^{-ki}$$

therefore

$$i = \frac{\ln(2)}{k}$$

$$\ln(2) \approx 0.693.$$

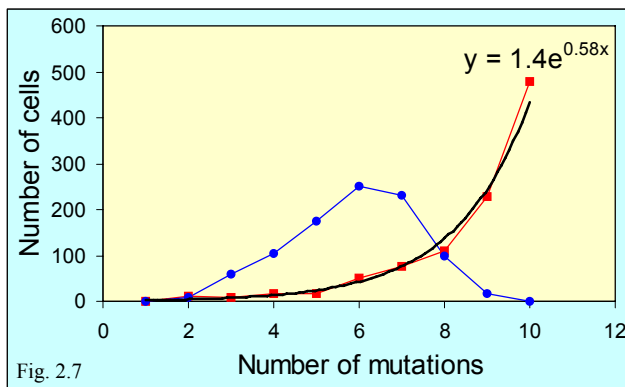
At the end, we consider a more complicated example. Assume a tissue which cells are randomly affected by mutations. Every cell genome has an equal chance to mutate. If we wait some time most cells will get more than one mutation. They start to accumulate. Now, we take a spreadsheet program and simulate the process. We compute 10 time intervals each for 1000 of our model cells simply by assigning a random number between 0 and 1 to each time interval. For random numbers above 0.5 a mutation strikes for values below 0.5 not. The Table below shows an example for 5 cells and two time intervals programmed with Excel. Now we plot the number of mutation events for each cell against the number of cells hit by this number. This is shown by the blue data points in the Figure below (Fig. 2.7). The result is quite complicated. Most often were 6 to 7 mutations, seldom 1 or 10 mutations per cell. However, in reality cells differ. They also differ in susceptibility to mutation events. We model this simply by assigning them random numbers that control how often they will be

A	B	C	D	E
1		Time intervals	Sum	Counter
2	Cell	1	2	=+LICZ.JEŽELI(D3:D7;2)
3	1	=+JEŽELI(LOS(<0.5;0;1)	=+JEŽELI(LOS(<0.5;0;1)	=+SUMA(B3:C3)
4	=+A3+1	=+JEŽELI(LOS(<0.5;0;1)	=+JEŽELI(LOS(<0.5;0;1)	=+SUMA(B4:C4)
5	=+A4+1	=+JEŽELI(LOS(<0.5;0;1)	=+JEŽELI(LOS(<0.5;0;1)	=+SUMA(B5:C5)
6	=+A5+1	=+JEŽELI(LOS(<0.5;0;1)	=+JEŽELI(LOS(<0.5;0;1)	=+SUMA(B6:C6)
7	=+A6+1	=+JEŽELI(LOS(<0.5;0;1)	=+JEŽELI(LOS(<0.5;0;1)	=+SUMA(B7:C7)

	A	B	C	D	E	F
1	Cell	Time intervals	2	Affectability	Sum	Counter
2	1	1	2			
3	1	=+JEŽELI(LOS()*\$D3<0.5;0;1)	=+JEŽELI(LOS()*\$D3<0.5;0;1)	+los()	=+SUMA(B3:C3)	=+LICZ.JEŽELI(E3:E7;2)
4	=+A3+1	=+JEŽELI(LOS()*\$D4<0.5;0;1)	=+JEŽELI(LOS()*\$D4<0.5;0;1)	+los()	=+SUMA(B4:C4)	
5	=+A4+1	=+JEŽELI(LOS()*\$D5<0.5;0;1)	=+JEŽELI(LOS()*\$D5<0.5;0;1)	+los()	=+SUMA(B5:C5)	
6	=+A5+1	=+JEŽELI(LOS()*\$D6<0.5;0;1)	=+JEŽELI(LOS()*\$D6<0.5;0;1)	+los()	=+SUMA(B6:C6)	
7	=+A6+1	=+JEŽELI(LOS()*\$D7<0.5;0;1)	=+JEŽELI(LOS()*\$D7<0.5;0;1)	+los()	=+SUMA(B7:C7)	

affected by a mutation. In essence, we model the process by a multiplication of two random numbers. Again we take 10 time intervals and look how many mutations each of the cells got. Suddenly the picture changed dramatically. Now, cell number and mutation number are related in a much simpler way, a way that can be described by an exponential function of the form $y = ae^{bx}$. Our spreadsheet program gives us automatically the associated parameters a and b, but it is easy to compute them manually by linearizing the equation using logarithms ($\ln(y) = \ln(a) + bx$).

This is of course a very simple example, but it shows us two important things. We modelled mutation events by assigning random numbers to assumed cells (simply the cells of the spreadsheet) and looked what



kind of distribution appears. This modelling approach is called a **Monte Carlo process**. We saw that identical cell properties resulted in a more complicated pattern than different ones. In other words, cell heterogeneity produced a pattern easier to analyse than a seemingly simpler homogeneous pattern. This is a very often found feature in nature. Heterogeneity, irregularity, or even chaos are not things that only complicate all. Very

often, they simplify and lead to more tractable relationships. This feature of heterogeneity makes it worth to study heterogeneity or even chaos in detail and the branch of mathematics dealing with this is the **chaos theory**.

3. Proportionalities

In an old German schoolbook I found the following exercise. 4 1/2 hens lay 15 3/4 eggs in 2 1/3 days. How many hens lay 27 2/5 eggs in 3 2/3 days? This is probably a typical every day problem in rural areas. So, how do peasants solve it? 15 3/4 eggs are laid in 2 1/3 days. Hence $15 \frac{3}{4} / 2 \frac{1}{3} = 6 \frac{3}{4}$ eggs per day per 4 1/2 hens, hence $6 \frac{3}{4} / 4 \frac{1}{2} = 1 \frac{1}{2}$ eggs per hen and day. We need 27 2/5 eggs in 3 2/3 days, this is $(27 \frac{2}{5} / 3 \frac{2}{3}) / 1 \frac{1}{2} = 5$ times the daily rate of a hen. We need 5 hens.

However, our solution is quite complicated and case specific. Using an appropriate mathematical formulation solutions of such ‘simple’ proportionalities are much easier. H hens lay E eggs in D days. Hence the daily rate is calculated to

$$\frac{E}{H * D} = \text{Eggs per hen and day}$$

This must hold for all other values of E, H, and D. Therefore

$$\frac{E_1}{H_1 * D_1} = \frac{E_2}{H_2 * D_2} \quad (3.1)$$

and

$$\frac{15.75}{4.5 * 2.333} = \frac{27.4}{x * 3.667} \rightarrow x = \frac{27.4 * 4.5 * 2.333}{15.75 * 3.667} = 5$$

In general solutions of tasks that deal with proportionalities can be solved by introducing proportionality equations that are constructed by referring to some standard value, in our case eggs per hen and days.

Let’s consider more realistic examples. Ideal gases can be described by kinetic gas theory. Robert Boyle (1627 - 1691) first realized that the volume of gases V is inversely proportional to the pressure upon it. We write

$$V \propto 1/P \text{ or } V P = \text{constant}$$

Joseph Gay-Lussac (1778—1850) found that under constant pressure a simple proportionality between volume and temperature T holds

$$V \propto T \text{ or } V / T = \text{constant.}$$

Now we have two simple equations that describe proportionalities. **Such equations can be combined by multiplying them.** Now we introduce a constant to transform into an ordinary equation

$$\frac{V P}{T} = \text{const}$$

Because gases are made of molecules or atoms we refer to a standard. We define **one mol as the weight of $N_A = 6.022169 * 10^{23}$ chemical units (atoms or molecules). N_A is the famous Avogadro constant.** We refer to one mol and assume a third proportionality. We assume that the volume needed by an ideal gas is proportional to the number of molecules or atoms it contains, hence to the number of mols. Denoting n for this number of mols we can now define a constant R

$$R = \frac{P V}{n T} \quad (3.2)$$

This is Lorenzo Avogadro's (1776 - 1858) law of ideal gases, a combination of simple proportionalities. R has a value of **8.3143 JK⁻¹mol⁻¹**. This has to be read as Joule per Kelvin and mol with one Joule being 1 Pascal*m³. The typical scientific notation for dimensions is the use of exponents

In chemistry we often need so-called **stoichiometric** calculations. For instance, Iron oxide has the general form Fe_xO_y. The analysis of such an oxide gave 40.95 g Iron and 15.6 g Oxygen. Determine the values of x and y. We need two things. First, the **mol weight** of iron (56 g / mol) and Oxygen (16 g / mol). O = 16, and C = 12. Second, we need a relationship that connects measured weights with mol weights. The quotient between the weights measured (having N molecules) should equal the quotient using only one molecule. The latter is provided by the reaction equation.

$$\frac{40.95}{15.6} = \frac{x * 56}{y * 16} \rightarrow \frac{x}{y} = \frac{3}{4}$$

This result can be generalized. Sugars have the general form H_xC_yO_z. Assume you got a probe that contains a grams hydrogen, b grams carbon, and c grams oxygen. We need again the mol weights (H = 1; C = 12; O = 16). Hence the proportion in the whole probe must be the same than in one molecule. Hence

$$a : b : c = 1x : 12y : 16z$$

$$\frac{a}{b} = \frac{x}{12y}, \frac{b}{c} = \frac{12y}{16z}, \frac{a}{c} = \frac{x}{16z}$$

Or even more general

$$a : b : c = xW_1 : yW_2 : zW_3$$

$$\frac{a}{b} = \frac{xW_1}{yW_2}, \frac{b}{c} = \frac{yW_2}{zW_3}, \frac{a}{c} = \frac{xW_1}{zW_3}$$

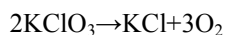
(3.3)

where W₁, W₂, and W₃ denote the mol weights of each element. Equation 3.3 can of course be used in the opposite direction. How many carbon dioxide originates if we burn 120 g carbon? The chemical equation is C+O₂→CO₂. Hence

$$\frac{120}{b} = \frac{12}{2 * 16} \rightarrow b = 320$$

We need 120 g carbon and 320 g oxygen to form 440 g CO₂.

The above equations however are valid only if we need one mol of each reactant to get one mol of the reagent. Consider the next reaction



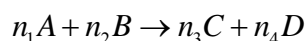
How many Kaliumchlorate do we need to get 637 ml Oxygen? Under normal conditions (273.15° K and 101325 Pa) **1 mol takes 22.4 litre** (V = 8,3143 * 273.15 / 101325 m³ = 22.4 l). Hence 637 ml oxygen are equivalent to 0.637 / 22.4 = 0.0284 mol. These are 0.0284*2*16 g. Now we apply equation 3.3. But we have to modify it. The left side gives the weights in the probe. This is the mol weight multiplied with the number of mols. Therefore

$$\frac{a(39.102 + 35.453 + 3 * 16)}{0.0284 * 2 * 16} = \frac{2 * (39.102 + 35.453 + 3 * 16)}{3 * 2 * 16} \rightarrow a = 0.01893$$

We need 0.01893*(39.102+35.453+3*16)g = 2.32g

At the end we get a general form of a **stoichiometric equation**. If we have a chemical reaction of the

form

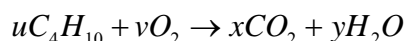


where A, B, C, and D are the reactants and n_1 to n_4 the number of atoms (or molecules or mols) we need, the relation between weights in the probe and mol weights of the reaction are given by

$$\frac{\text{weight } A}{\text{weight } C}(\text{probe}) = \frac{n_1 \text{Molweight}(A)}{n_2 \text{Molweight}(C)}(\text{probe}) = \frac{n_1 \text{Molweight}(A)}{n_2 \text{Molweight}(C)}(\text{chemical equation}) \quad (3.4)$$

Identical equations hold for the other combinations of reactants.

In the simplest case all reactants are known. For instance Butane reacts with oxygen to form carbon dioxide and water. First establish the reaction equation. The general form of this equation is



We need u, v, x, and y. This can be done by applying the simple scheme of the great German mathematician Carl Friedrich Gauß (princeps mathematicorum, 1777 - 1855).

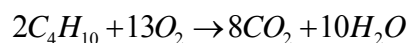
$$4u = x$$

$$10u = 2y$$

$$2v = 2x + y$$

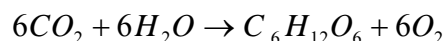
Now we have three possibilities. We can try to solve this system by hand. In this case this is simple but in other cases this might be quite complicated. We can use the internet for solving. ScienceSoft (www.sciencesoft.at/) provides a nice small internet service for solving stoichiometric equations. We can also apply a math program. A Mathematica solution looks as follows.

Now we have to seek for the smallest y giving natural u, v, and x. This is apparently $y = 10$. u becomes 2, v, 13, and x 8. Therefore



But now compare our result with that of the internet service. In our case you count on both sides 26 O, ScienceSoft computes 12 O on the left but 13 O on the right side. This result is obviously erroneous. Don't trust every result of a seemingly well introduced and nice looking program!

Let's consider a more complicated example. How many carbon dioxide do you need to get 50 g glucose. First, you need the reaction formula.



Hence

$$\frac{x}{50} = \frac{6(12 + 2*16)}{6*12 + 12*1 + 6*16} \rightarrow x = 73.33$$

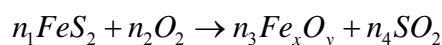
By the same logic we get the amount of water and oxygen

$$\frac{y}{50} = \frac{6(2 \cdot 1 + 16)}{6 \cdot 12 + 12 \cdot 1 + 6 \cdot 16} \rightarrow y = 30$$

$$\frac{z}{50} = \frac{6 \cdot 2 \cdot 16}{6 \cdot 12 + 12 \cdot 1 + 6 \cdot 16} \rightarrow z = 53.33$$

Of course, the amounts of reactants of both side of the chemical equation are the same: $73.33\text{g} + 30\text{g} = 50\text{g} + 53.33\text{g} = 103.33\text{g}$.

Equation 3.2 is a general description of a stoichiometric reaction between two reactants. It describes the proportionalities with regard to weight. But it can also be used to infer the chemical equations. For instance the oxidation of 20g iron sulfide (FeS_2) gave 21.36g SO_2 and an unknown iron oxide. We needed 10.28 l oxygen. Determine the chemical formula of this oxide and the reaction equation. First, the reaction equation must have the general form



Second one mol oxygen = 32 g takes under normal conditions 22.4 l. Hence 10.28 l are 0.459 mol. These are 14.685g. Let the weight of Iron oxide be IO. The sum of weights at both sides must be equal. $20 + 14.685 = \text{IO} + 21.36$. Hence $\text{IO} = 13.325\text{g}$.

Now we apply equation 3.2

$$\frac{20}{21.36} = \frac{n_1(55.85 + 2 \cdot 32)}{n_4(32 + 2 \cdot 16)} \rightarrow \frac{n_1}{n_4} = \frac{1}{2}$$

$$\frac{20}{13.325} = \frac{n_1(55.85 + 2 \cdot 32)}{n_3(x \cdot 55.85 + y \cdot 16)} \rightarrow \frac{n_1}{n_3} = 0.0125(55.85x + 16y)$$

$$\frac{20}{14.685} = \frac{n_1(55.85 + 2 \cdot 32)}{n_2(2 \cdot 16)} \rightarrow \frac{n_1}{n_2} = 0.364$$

$$\frac{21.36}{13.325} = \frac{n_4(32 + 2 \cdot 16)}{n_3(x \cdot 55.85 + y \cdot 16)} \rightarrow \frac{n_3}{n_4} = \frac{39.925}{(55.85x + 16y)}$$

$$\frac{21.36}{14.685} = \frac{n_4(32 + 2 \cdot 16)}{n_2(2 \cdot 16)} \rightarrow \frac{n_2}{n_4} = 1.375$$

$$\frac{13.325}{14.685} = \frac{n_3(x \cdot 55.85 + y \cdot 16)}{n_2(2 \cdot 16)} \rightarrow \frac{n_2}{n_3} = 0.0344(55.85x + 16y)$$

We have six equations and seven unknown variables. But we know even more. We apply again the Gauß scheme

$$n_1 = n_3x$$

$$2n_1 = n_4$$

$$2n_2 = n_3y + 2n_4$$

Now we have nine linear equations with six unknown variables. Such a system can be solved. From equation 1, 3 and 5 we infer that $n_1:n_2:n_4 = 1:2.75:2$. The smallest natural numbers solving these relations are $n_1 = 4$, $n_2 = 11$, and $n_4 = 8$. Now it is easy to solve the other

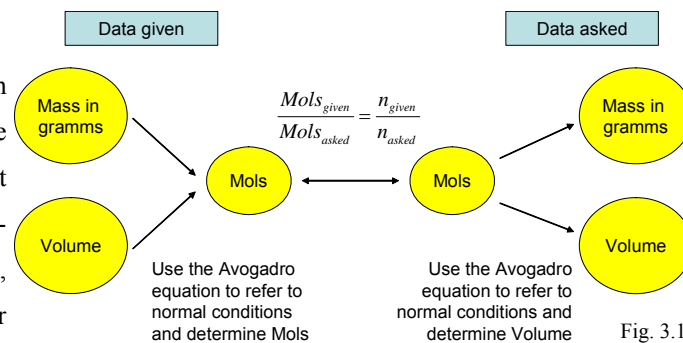
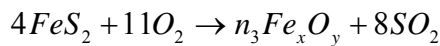


Fig. 3.1

equations for x and y . We get



We took 22 O and got 16 O in SO_2 . Hence $n_3 * y = 6$. Additionally, $n_3 * x = 4$. The smallest three natural numbers fulfilling this equation are $n_3 = 2$, $x = 2$, and $y = 3$. At the end we get

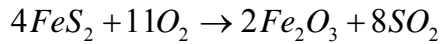


Fig. 3.1 gives a general scheme how to solve stoichiometric equations. First, you have to refer to mols as the most suitable measure. Then you apply a proportional equation (Equation 3.2) in which n stands for the coefficient.

Now we consider another form of a chemical proportion equation.

You need 50ml of 70% solution of ethanol in water. You have 90% ethanol. What to do? This is also a problem that is best solved using a systematic approach. You need a series of linear equations containing volumes (or weighs) and concentrations. The first equation contain volumes

$$x \text{ (ml water)} + y \text{ (ml ethanol)} = 50 \text{ (ml solution)}$$

The second equation gives concentrations. The total amount of ethanol remains constant. Hence

$$0.9 y \text{ (ml ethanol)} = 0.7 * 50 \text{ (ml solution)}$$

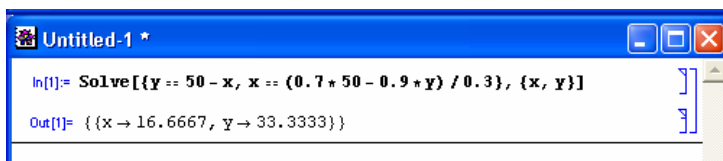
This second equation contains 1.0, 0.9, and 0.7. Mathematicians avoid using percents. Instead it is most often easier to use a range from 0 to 1. Hence 1.0 stands for 100%, 0.7 for 70%, and 0.9 for 90%. Both equations can easily be solved

$$y = 0.7 * 50 / 0.9 = 38.89 \text{ ml 90\% ethanol. Therefore } x = 50 - 38.89 = 11.11 \text{ ml water.}$$

The same task in a different way. You need 50 ml 70% ethanol solution. You have solutions of 30% and 90%. Again

$$x \text{ (ml 30\%)} + y \text{ (ml 90\%)} = 50 \text{ (ml 70\%)} \quad \text{and now}$$

$$0.3x \text{ (ml 30\%)} + 0.9y \text{ (ml 90\%)} = 0.7 * 50 \text{ (ml 70\%)}$$



Hence $x = 16.67$ ml and $y = 33.33$ ml.

Note that standard math programs are able to solve algebraic equations like above.

Such a solution looks as follows. Note that in the example before with six equation it was not possible to apply the math program (at least not in a simple way) because we had proportional equations and the simplest solution the program proposes is the trivial solution that all variables n equal zero.

An example of a **nonlinear proportionality** is the **Weber-Fechner law** of physiology. The perception intensity P of a stimulus is proportional to the logarithm of the physical intensity I of the stimulus

$$P = k \ln\left(\frac{I}{I_0}\right) + c$$

(3.5)

where I_0 is most often an arbitrary set baseline and c a constant.

Hence, signal perception increases with the logarithm of signal intensity. Such a law holds for instance for photoreception, acoustics or olfactorical stimuli.

Of major importance is the Weber Fechner law in acoustics. Perceived loudness L follows equation 3.5 and is defined in db (DeziBel) by the following equation

$$L_x = 10 \log_{10} \frac{I}{I_0} = 20 \log_{10} \frac{p_x}{p_0} [dB] \tag{3.6}$$

p_0 is an arbitrarily predefined lower sound pressure level of $2 \cdot 10^{-5}$ Pascal. p_x is the sound pressure we measured. For instance a sound pressure of $4 \cdot 10^{-2} \text{ Nm}^{-2}$ is equivalent to $L = 20 \cdot \log_{10} (2 \cdot 10^3) = 66 \text{ dB}$. With equation 3.4 we can compare different sound pressures and compute the difference in loudness we perceive. Therefore, if we compare two sources of loudness the difference ΔL is

$$L_1 - L_2 = \Delta L = 10 \log_{10} \left(\frac{I_1}{I_0} \right) - 10 \log_{10} \left(\frac{I_2}{I_0} \right) = 10 \log_{10} \left(\frac{I_1}{I_2} \right) [dB]$$

Hence, a hundredfold increase in intensity causes an increase of 20dB.

Signal transmission and regulation

In physiology we deal with signal transmission and regulation. Examples are the induction or inhibition of biochemical processes, thermoregulation, secretion, gene activation, or enzyme regulation. The mathematical theory of regulation (the science of cybernetic) is very complex, but is based on some very simple assumptions. Fig. 3.1 shows a simple regulatory system. The demand value (fixed) is compared with an actual value (measured by

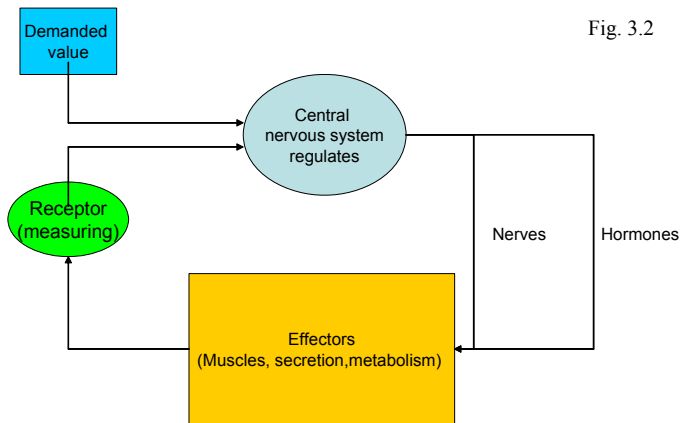


Fig. 3.2

receptors). The is done by the central nervous system. The difference between both values is the output signal to the effectors. If this difference is zero, no further change of the effectors state is needed. Hence, the basic mechanism of regulation is a difference between demand and actual value. These values can be measured and incorporated into a model of the regulatory system under study.

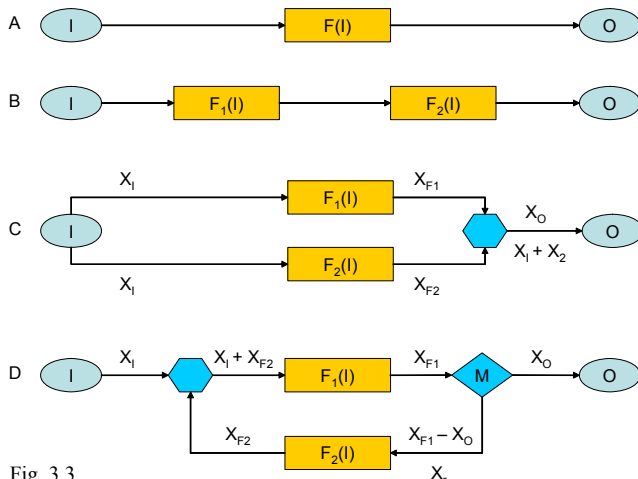


Fig. 3.3

Fig. 3.2 shows four basic models of signal transmitting and regulation. We have an input signal I that is handled by a function F to give the output signal O . Mathematically $O \propto F(I)$. In B we have two handling operations F_1 and F_2 that are in line. O is now proportional to both functions. Hence $O \propto F_1(I) F_2(I)$.

In C we have again two function that operate in parallel. The input signal X_1 goes to F_1 and F_2 . O is the output of the sum of F_1 and F_2 . Hence

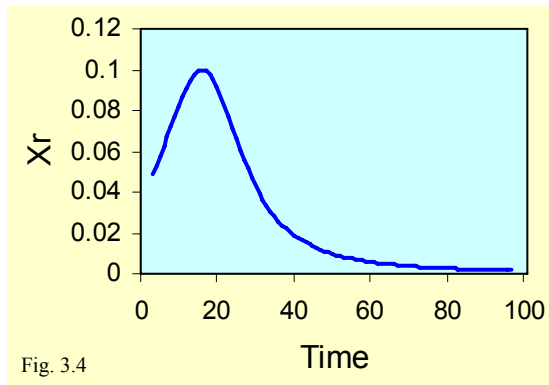


Fig. 3.4

$$O \propto F_1(I) + F_2(I).$$

In D the output signal is regulated. X_{F1} is measured by M. M gives a signal of value $O - X_{F1}$ to F_2 that produces the output signal X_{F2} . This adds to X_I . Hence if X_{F1} is too small X_{F2} acts to enhance it, if it is too large, X_{F2} dampens it. Mathematically: After regulation X_{F1} should equal X_O

$$X_O = F_1(X_I + X_{F2}) - X_r$$

$$X_{F2} = F_2(X_{F1} - X_O).$$

We solve these equations for X_O and get

$$X_O = \frac{F_1}{1 + F_1 F_2} X_I + \frac{F_1 F_2 - x_r}{1 + F_1 F_2}$$

At perfect regulation x_r becomes zero and the last term is a constant. Hence

$$X_O = \frac{F_1}{1 + F_1 F_2} X_I + c \tag{3.7}$$

This is the signal equation of a simple self regulatory system.

We can now plot X_r against time (the number of cycles) after an initial disturbance of ΔI . This is shown in Figure 3.4. After an initial peak X_r slowly returns to zero that mean X_{F1} again equals X_O .

How much information is transmitted by the output signal? To answer this question we have to define what we mean by information. The whole output signal O contains two elements

$$O = H + R \tag{3.8}$$

where H is called the **entropy** of a system, the amount of information it contains. R denotes the **redundancy**, the amount of **noise** the output contains. Now suppose a digital signal transmission. The maximum value of H is the number of interpretable (0,1) switches. The maximum information content is then defined as the binary logarithm (logarithm to base 2) of the number of switches n .

$$H_{\max} = lb(n) \tag{3.9}$$

To define the meaning of information Claude Shannon (the American founder of information theory, 1903-1969) referred in 1949 to the concept of chemical entropy S as derived by the German chemist Ludwig Boltzmann (1844 - 1906). Assume a state of two atoms (or signs) take a certain state with probability p_1 and p_2 . The probability that both atoms have this state $S(p_1 p_2)$ is then $p_1 * p_2$ (this is the multiplication law of probabilities we will deal with in the second part of this lecture). Now assume that entropy (a function of probability) is an additive concept. The total amount of entropy is the sum of all entropies of the atoms. Hence

$$S(p) = S(p_1) + S(p_2)$$

and

$$S(p_1 p_2) = S(p_1) * S(p_2)$$

Both equations only have a solution if entropy is a logarithmic function of probability.

$$S(p) = a * \log(p_1 p_2) = a \log(p_1) + a \log(p_2) = S(p_1) + S(p_2).$$

Boltzmann used this concept, referred to natural logarithms and defined the entropy by a fundamental equation

$$S = k \ln(p) \quad (3.10)$$

The entropy is proportional to the natural logarithm of the probability of its occurrence. k is the Boltzmann constant and equals R / N_A . p is a measure of the probability of occurrence. The lower this probability the more unstructured is the system. The Austrian chemist Erwin Schrödinger (1887 - 1961) defined the inverse of p ($D = 1 / p$) as a **measure of uncertainty or irregularity** and got

$$S = k \ln(p) = k \ln(1 / D) = -k \ln(D)$$

Shannon, dealing with digital signal transmission now referred to binary logarithms and defined the **information content of a signal** as

$$H_i = -\text{lb}(p_i) \quad (3.11)$$

Where p_i denotes the frequency of occurrence (its probability) of a signal i . H is measured in bits (binary information units needed to define a signal). The information content of this particle with respect to others is H_i multiplied with its relative frequency. However, this is nothing more than its probability to occur, hence p_i . If we deal with a whole system of n atoms (elements) we get the mean information content of a signal. This is the sum of all entropies.

$$H = -\sum_{i=1}^n p_i \text{lb}(p_i) \quad (3.12)$$

This is the most often used definition of **information**. H is generally used as a **measure of diversity** in its broadest sense. It treats information solely under a concept of probability of occurrence. It does not refer to any content. Hence it does not discriminate between information that has sense and information without any sense.

If we have two signals we can measure information from the divergence of signal A from signal B. If signal B is a theoretical expectation, the information content of signal A is given by the divergence of A from B. This measure of information is known as the **Kullback-Leibler information** and finds much application in biological modelling. It is defined as

$$H = -\sum_{i=1}^n p(A_i) \ln \frac{p(A_i)}{p(B_i)} \quad (3.13)$$

The maximum information content is reached if all signals have equal probability of occurrence. Hence $H = -n * p * \text{lb}(p) = n * 1 / n * \text{lb}(1 / n) = \text{lb}(n)$.

From this we get an additional measure how even the information content is distributed among the signals

$$E = \frac{H}{\text{lb}(n)} \quad (3.13)$$

This measure is called **evenness**.

4. Some basic functions and their application in biology

What is a function?

This fourth lecture deals with several important functions; functions that are often used in many branches of biology.

First of all, what is a function? A function is a mathematical term that combines two or more variables in one or more equations. Normally, a function contains one so-called **dependent variable y** and a set of **independent variables x_i**. Typical functions with only two variables are denoted by $y = f(x)$. Examples are

$$y = ax + b$$

$$y = \ln(x) + b$$

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 \dots = \sum_{i=0}^n a_i x^i$$

In the latter example we used a summation sign to shorten the function, the large Greek Sigma = Σ . The first of these examples is of course a linear and the second a logarithmic relationship. Especially important is the last example. This is the general form of a so-called **algebraic function**. This is a function with only one independent variable x that is set to powers from 0 to n. The constants a_i are most often real numbers.

A special feature of a true function is that the set of independent variables x_i determine the value of y in such a way, that for any set of values x_i there is one and only one value y. The above three cases fulfil this prerequisite. However, there are examples where for one x-value several y-values exist. For instance, if we want to describe a circle (Fig. 4.1) we have to use the following equation

$$x^2 + y^2 = r^2$$

that follows immediately from the law of Pythagoras. Fig. 4.1 shows us that now every value of x gives two values of y. These are

$$y_1 = +\sqrt{r^2 - x^2} \text{ and } y_2 = -\sqrt{r^2 - x^2}$$

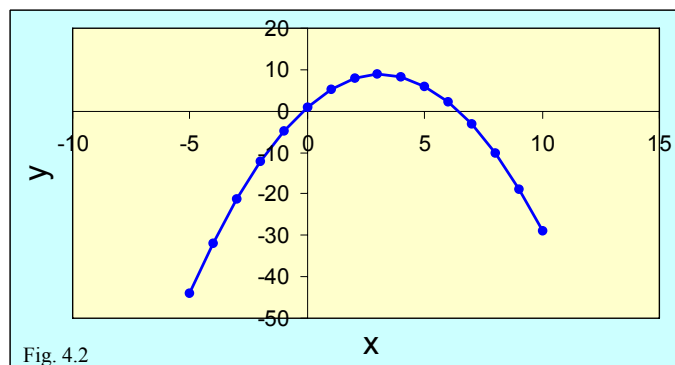
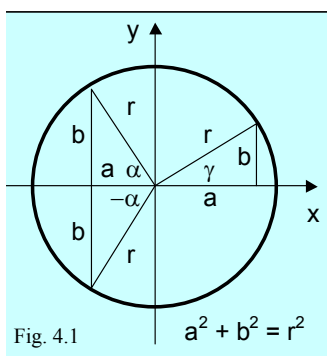
In this case we speak of **relations**.

Fig. 2.2 shows a plot of the algebraic function

$$y = 5x - 0.8x^2 + 1$$

in the boundary $-5 < x < 10$.

We detect two important points where $y = 0$, the so-called **roots** or zeroes of the function. Computing them is easy. We use the binomial relationship



$$(a+b)^2 = a^2 + 2ab + b^2$$

$$-0.8x^2 + 5x + 1 = 0$$

$$\downarrow$$

$$x^2 - \frac{5}{0.8}x + \left(\frac{1}{2 \cdot 0.8}\right)^2 = \left(\frac{1}{2 \cdot 0.8}\right)^2 + \frac{1}{0.8}$$

$$\downarrow$$

$$\left(x - \frac{5}{1.6}\right)^2 = \left(\frac{5}{1.6}\right)^2 + \frac{1}{0.8} \rightarrow x = \pm \sqrt{\left(\frac{5}{1.6}\right)^2 + \frac{1}{0.8}} + \frac{5}{1.6}$$

$$x_1 \approx 6.44; \quad x_2 \approx -0.19$$

Below the solution of a math program is shown.

```

Untitled-1 *
In[1]:= Solve[0 == -0.8 * x^2 + 5 * x + 1, x]
Out[1]:= {{x -> -0.19398}, {x -> 6.44398}}
    
```

In general, algebraic functions have analytical solutions only up to the fourth order (that means up to the fourth power), although these are quite complicated. For

functions of higher order most often only approximate solutions exist.

Functions are characterized by ranges over which the function is defined. This **range** (or **domain**) is an ordered set of the y and x-variables that defines where the function is valid. Often, functions have boundaries. Consider the logarithmic function $y = \ln(x)$. This function is not defined for values of x less than or equal 0. In this case we denote

$$y = \ln(x); \text{ for } x > 0$$

More important are cases where the y-values are bounded. Look at the function $y = 1/x + b$. If x goes to infinity $1/x$ goes to 0 and y to b. Mathematically speaking, **y asymptotically reaches b**. But y will never be equal to b. Below we will discuss this point in more detail.

Important are shifts of functions along the x- and y-axes. Fig. 4.3 shows the function $y = x^2 - x + 1$.

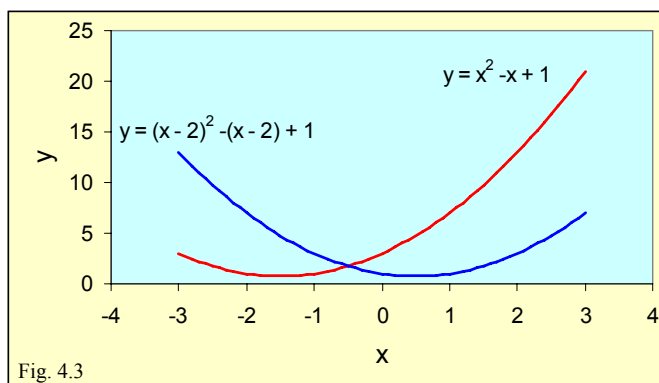


Fig. 4.3

Using $x - 2$ instead of x shifts the whole function two units to the right without changing the shape of the function. This is an important feature. In general the transformation

$$y = f(x) \rightarrow y - y_0 = f(x - x_0)$$

shifts a function x_0 units to the right and y_0 units up.

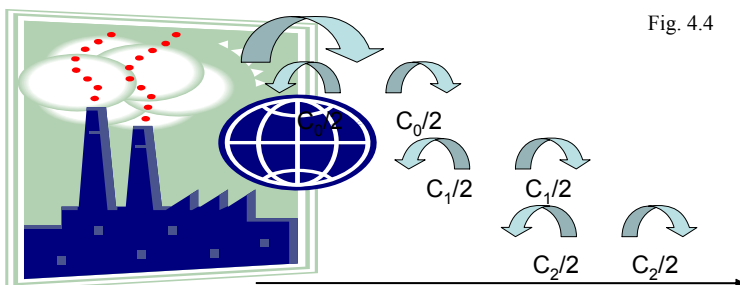


Fig. 4.4

The exponential function

Look at Fig. 4.4. A industrial plant produces some sort of air pollution. If we assume a day without wind this pollution has its highest concentration near the plant. The farther away we are the lower the concentration

will be. We can model this process by a simple assumption. We assume that the particles have at every distance an equal chance to be drifted either to the left or to the right side as shown in the Figure. The concentration after an imaginary first step is $1/2C_0$, with C^0 being the initial concentration at the chimney. This process is identical at every step. Hence at step two the concentration is $1/2(1/2C_0)$. At step x we get

$$C_x = C_0 \left(\frac{1}{2}\right)^x = C_0 e^{-\ln(2)x} = C_0 e^{-0.693x}$$

This is the first biologically important function we consider, the **exponential function**. It has the general form

$$y = ae^{kx}$$

where a and k are constants. One of the most important processes of this type is the radioactive decay. Assume a fraction of p atoms decaying in a certain time range t_1 . $1-p$ atoms remain. This process is absolutely independent of the initial number of atoms and the time already spent for decaying. Therefore, the fraction of atoms N_x that decayed after x time windows is

$$N_x = N_0 p^x = N_0 e^{\ln(p)x} = N_0 e^{ax}$$

Because p was the fraction that decayed it is a value between 0 and 1. The logarithm of p is therefore negative. We can now define a value $k = -a = -\ln(p)$ and get the general relationship

$$N_x = N_0 e^{-kx} \tag{4.1}$$

We heard already in the previous lecture that this exponential function is a general description of the number of entities having a certain property in a **multiplicative process**. It refers to such different processes like radioactive decay, species extinction, substrate concentrations, or diffusion. N_x has a lower boundary of 0. In mathematical language we denote $N \rightarrow 0$.

The number of entities not having this property is of course

$$N_{-x} = N_0 (1 - e^{-kx}) \tag{4.2}$$

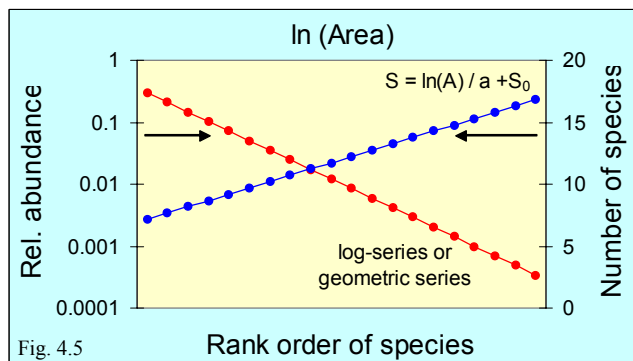
In our case, this would be the number of atoms that did not decay.

Of especial importance is again the value of x where N_x is exactly $N_0 / 2$, the half time. We saw already that this value is $N_{1/2} = \ln(2)/k$.

The logarithmic function

Let's assume a community of plant species that live together in a certain area A . The most abundant species occupies a fraction of p of this area. The second abundant species takes again a fraction of p of the remaining area and so on. This is again a process that can be described by an exponential function (why?) and the abundance of the i -th species is

$$N_i = N_0 e^{-ai} \tag{4.3}$$



In this function N_0 is the total abundance (the total number of individuals) of all species combined. Rearranging this equation gives

$$\ln \left(\frac{N_i}{N_0} \right) = -a i$$

The term N_i / N_0 is called the **relative abundance** of a species i . If we plot the relative abundances on a logarithmic scale against the species number i we get therefore a straight line as shown in Figure 4.5. Such a relationship is called a **geometric series**. More than 60 years ago the Japanese ecologist I. Motomura proposed it for the first time to describe relative abundances of animals.

Now we might be interested in the question how many species we would find in a certain part A_i of the total area A in which our community lives. If the least abundant species has a density of 0.01 per unit area, we would expect one individual and therefore also this species to be found at 100 units of area. For a density of 0.0001 we already need an area of 10000 units of area. In general, the area needed to find a species is the reciprocal of its density. It follows therefore

$$-\ln(A_i) = -a i + \ln(A_0)$$

and

$$i = \frac{1}{a} \ln(A_i) + \frac{1}{a} \ln(A_0)$$

This is a general form of a logarithmic function $y = a \ln(x) + b$. We find the constant b if we set A_i to one, in other words we consider the number of species i at one unit of area. This number is $\ln(A_0) / a$ and is frequently denoted as S_0 . A plot of i against $\ln(A)$ should give a straight line with the slope $1/a$. Setting $i = S$ we have

$$S = \frac{1}{a} \ln(A) + S_0 \tag{4.4}$$

We got an important relationship. If the relative abundances of the species in a community follow a geometric series (or the very similar **log-series**) the associated **species – area relationship** should follow a logarithmic function. This type of **species—area relationship** was once proposed by the American botanist Henry Gleason (1882 - 1975) and is of major importance in plant ecology as a tool for estimating species numbers at the local scale. Logarithmic functions are of course only defined for values of x above 0.

The power function

Surely the most important function in biology is the **power function**. In morphology these are often called **allometric functions**. Look at the following Figure 4.6. Given are mean body weights of some mammal species and their respective brain weights. A plot of brain weight against body weight on a double log scale (both the x- and the y-axis are log transformed) results in a straight line where the data points scatter around an as-

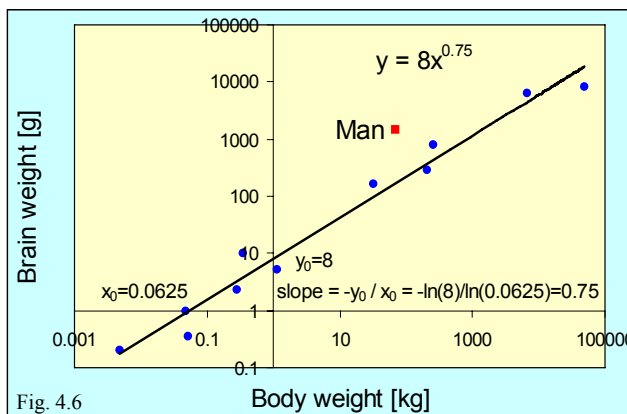


Fig. 4.6

sumed so-called **regression line**, a curve that is chosen in such a way that the data points are nearest to this curve. Only Man is outstanding with its high brain weight. Our spreadsheet programs automatically compute such curves from predefined functions and give these functions. In the statistics lecture we will see how to compute such functions and what has to be considered when using them.

We might therefore describe the dependence of brain weight on body weight in mammals by the following function

$$\begin{aligned} \ln y &= a \ln(x) + \ln(b) \\ \text{or} \\ y &= bx^a \end{aligned} \tag{4.5}$$

This is the general form of a power or allometric function. a is the slope of the function, often termed the **scaling factor**. b gives the y -value at $x = 1$. $\ln(b)$ is therefore the **intercept** on a double log scale. For $a = 1$ we get $y = bx$ and this form of a linear dependence (simple proportionality) appears to be a special form of a power function (where the function goes through the origin). However, we immediately notice an important difference between a linear dependence and a power function. The latter is defined only for values x above 0.

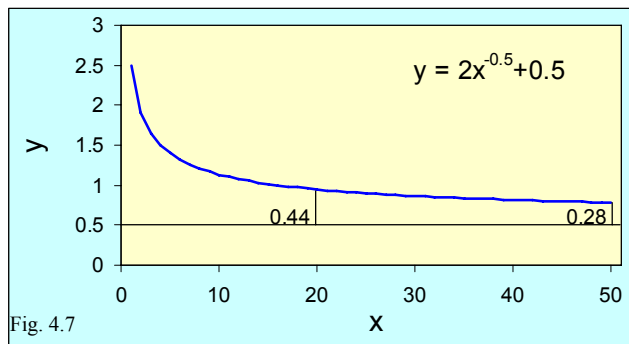


Fig. 4.7

The next Figure 4.7 shows a power function with a negative slope. These special types of power functions are often called **hyperbola**. We notice that this function has a noticeable curvature only at small values of x . It seems that for $x \approx 20$ something like a threshold appears. For x values larger than about 20 the y -values are very small in relation to the initial ones. The absolute difference between y_{50} and y_{20} is only 0.16, whereas the difference between y_{20} and y_1 is 2.06. Above $x \approx 20$ the curve appears to be nearly straight. If, for instance, this would be a plot of reproduction rate (the total number of offspring of a species) against body weight (a relation that is well described by a power function) we would conclude, that for larger animals the above relationship is of little value. At higher body weights total offspring number is roughly constant and relatively small. Such a pattern is quite often found in nature. We speak of a **heavy tail** or a **heavy tail pattern**.

Power functions are of major importance in biology and we have to know about some important features. One important feature is that the slope of a power function is independent of the way we measured x and y . For instance in our example above, we measured body weight in kg and brain weight in g. What is if we change into grams? We get

$$B[g] = aW^z[kg] \rightarrow B[g] = a(1000W)^z[g] = a(1000^z)W^z[g] = kW^z[g]$$

where k contains the constants and is the new intercept. In other words, the intercept of a power function contains the units of measurements, but the exponential term is independent of how we measured our data. Additionally, constant a and slope z should be independent. But be careful, this relationship is only valid if we deal with an infinite universe. But most often, we have data from experiments or observations. In these cases, the data points have a limited range of allowed values. Now a and z are not longer independent and it depends on the data structure how they are related.

Above, we heard about relative abundance or species—rank order distributions. We dealt with a log-series as an important example for this. What is if the species rank order can be described by a power function?

$$N_i = N_0 i^{-a} \rightarrow \ln\left(\frac{N_i}{N_0}\right) = -a \ln(i)$$

Again, the area needed to find a species is the reciprocal of its abundance. It follows therefore

$$\ln(i) = -\frac{1}{a} \ln\left(\frac{A_0}{A_i}\right) \rightarrow i = A_0^{-1/a} A_i^{1/a}$$

In this case we get an allometric species—area relationship that is generally written as

$$S = S_0 A^z \tag{4.6}$$

This is the most often found type of species—area relationship and frequently serves as a general starting point for ecological modelling.

It is easy to combine allometric functions to reach in new dependencies. Consider the next example. 60 years ago Max Kleiber found that metabolic rates of animals scale to body weight to the power of 0.75 .

$$M \propto W^{0.75}$$

We use this form to describe a power function because the intercept is for us of minor importance. Body weight scales to body length to the power of about 2.6 (see equation 2.4 above) to 3. The latter value 3 is of course the theoretical (geometric) expectation. Therefore

$$W \propto L^{2.6 \text{ to } 3.0}$$

and

$$M \propto L^{0.75 * (2.5 \text{ to } 3.0)} \propto L^{1.875 \text{ to } 2.25}$$

Additionally, body weight and mean abundance of species are connected by a power function with a negative slope ($D \propto W^{-z}$). We now simply multiply both scaling laws and get

$$MD \propto W^{0.75} W^{-z} \propto W^{0.75-z} \propto L^{2-z} \tag{4.7}$$

In 4.7 I rounded 1.875 to 2.25 to a value of 2. Now, we don't deal with an individual species but with all individuals of a species, the population. The product MD is the total metabolic rate of all species of the community considered.

For mammals, the British ecologist J. Damuth found that z takes values between 0.5 and about 1 with a mean close to 0.75. Hence he got

$$MD = B \propto W^{0.75-0.75} = \text{const} \tag{4.8}$$

We conclude that total metabolism (often approximated by population biomass, the sum of all body weights of all organisms in a population of a given species) is roughly constant for species with different body weight. We obtained this result by a simple combination of two independent scaling laws. This is known as the **equal biomass** or **energy equivalence rule**, firstly formulated in

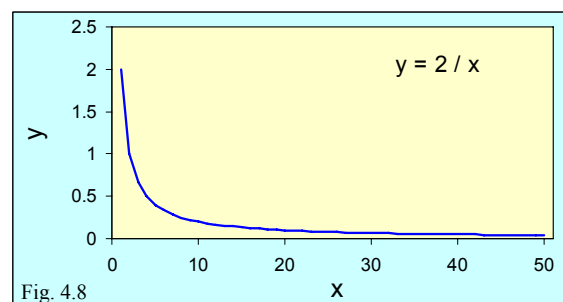


Fig. 4.8

1981 by Damuth. However, if z differs from 0.75, total population biomass, and therefore population energy turnover, is a function of mean individual (species) body weight. For arthropods, for instance, z -values > 1 have been reported. That means that smaller organisms would have a much higher total metabolic rate (energy turnover) than larger ones.

The hyperbolic function

Consider a power function with a slope of -1 . That means

$$\begin{aligned}
 &y = ax^{-1} \\
 &\text{or} \\
 &xy = a = \text{constant}
 \end{aligned}
 \tag{4.9}$$

This is a function that is shown in Fig. 4.8. It is called a hyperbola. For instance, reproductive output (the number of offspring) scales to body weight at $R \propto W^{-x}$. But at the same time reproductive costs (the energy needed to produce offspring) scales to body weight at $C \propto W^z$. The product of costs and reproductive output should therefore be

$$CR \propto W^{z-x}$$

If we now compare different species of animals of the same mean body weight we expect that

$$CR = \text{const}$$

For species of similar body weight the relation between C and R should therefore be a hyperbola. We notice that a hyperbola is symmetrical to both axes and has no intercepts. Sometimes the smallest distance from the origin is of interest. From the symmetry condition follows that this smallest distance is exactly $d_{\min} = \sqrt{2a}$.

The inverse hyperbola

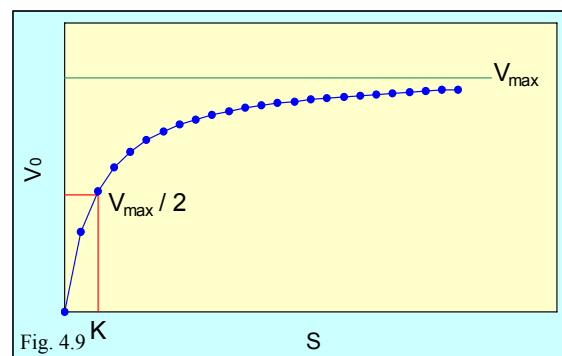
More important than the true hyperbola is the **inverse hyperbola**. Well known is this function in the form of the so-called **Michaelis – Menten equation**.

Consider an enzyme E that binds to a substrate S : $E + S \leftrightarrow ES$. Now consider that moment where this process reaches a **stationary or equilibrium state**. Let E_0 be the initial enzyme concentration. Then, the speed to produce ES V_1 will equal the decay rate of ES V_0 through dissociation. V_1 is approximately proportional to the substrate concentration and to the concentration of E . The latter is $(E_0 - ES)$. This gives the following equation

$$k_1 S (E_0 - ES) = k_2 ES$$

Now we solve this equation for ES and get

$$ES = \frac{k_1 E_0 S}{k_2 + k_1 S}$$



The speed of the enzymatic reaction is highest at the initial enzyme concentration E_0 : $V_{\max} \propto E_0$. Additionally we divide the above equation through k_1 and subsume the constants k_1 ad k_2 in one constant K . We get

$$V_0 = \frac{V_{\max} S}{K + S} \quad (4.10)$$

This is the Michaelis-Menten equation that describes the kinetics of an enzyme. If we plot V_0 against the substrate concentration S a hyperbola appears as shown in Fig. 4.9. This hyperbola is limited. Asymptotically, it approaches (but never reaches) a maximum value. This maximum is of course V_{\max} . In other words

$$\lim_{S \rightarrow \infty} V_0 = V_{\max}$$

Another interesting value to describe the whole process is K . Assume a reaction speed V_0 that is half of V_{\max} . We get

$$\frac{V_{\max}}{2} = \frac{V_{\max} S}{K + S} \rightarrow K + S = 2S \rightarrow K = S$$

K is exactly that concentration of S where $V_0 = V_{\max} / 2$.

The Michaelis-Menten equation is one very important example for a whole class of functions (the Monod functions) described by the equation

$$y = \frac{af(x)}{b + f(x)} \quad (4.11)$$

They all have their limes at a . The equation may also be linearized. We take the inverse of both variables x and y and get

$$\frac{1}{y} = \frac{b}{a} \frac{1}{f(x)} + \frac{1}{a}$$

A plot of $1/y$ against $1/f(x)$ is therefore a straight line with slope b/a and intercept $1/a$. This method is called the **Lineweaver - Burk transformation**.

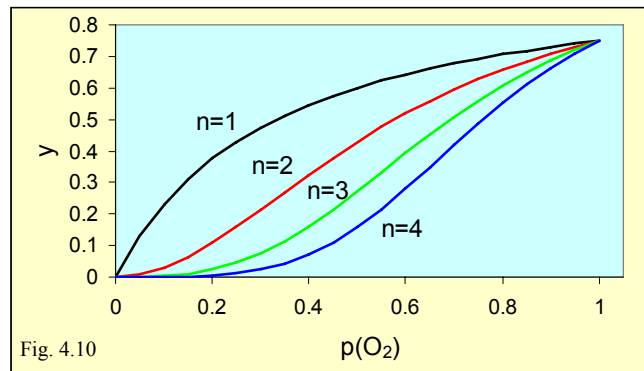
An extension of the hyperbola model has other applications. For instance haemoglobin and myoglobin bind oxygen according to the partial pressure of O_2 . According to the law of masses there exists an equilibrium between, for instance, the concentrations of myoglobin (M) and oxygen (O) on one side and the MbO complex on the other side

$$K = \frac{[MbO]}{[Mb][O]}$$

Denoting y for $[MbO]$, $p(O_2)$ for the partial pressure of oxygen and using $[MbO] + [Mb] = \text{const}$ we get

$$K = \frac{y}{(\text{const} - y)p(O_2)} \rightarrow y = \frac{K p(O_2)}{\text{const} + K p(O_2)}$$

Myoglobin is monomer. For di-, tri-, or tetramer Molecules we have to modify the latter equation and get the so-called **Hill equation** of oxygen binding



$$K = \frac{y}{(const - y)p(O_2)^n} \rightarrow y = \frac{K p(O_2)^n}{const + K p(O_2)^n}$$

that describes the concentration of the MbO complex in dependence of oxygen partial pressure. Fig. 4.10 shows plots of this equation for various n . For $n=1$ the Hill equation is equivalent to the Michaelis - Menten model of enzyme kinetics. Indeed both processes are functionally similar because they describe the binding of a substrate to a protein. For $n > 1$ the plots approach a **sigmoidal shape**. The Hill equation can be seen as an extension of the Michaelis - Menten model.

5. Handling a changing world

One of the most often met mathematical problems in biology is to measure changes for instance in time or in nutrient concentration or in blood pressure or or....

Consider a car that drives with a constant speed of 50 km / hour. We may formalize this by writing

$$f(t+\Delta t)=f(t)+v\Delta t$$

In other words, the distance from the starting point $f(t)$ reached after time $t+\Delta t$ is the distance after time t plus the way driven during the time interval Δt . The latter is the product of speed v and Δt . We can rearrange the above equation and get

$$\frac{f(t+\Delta t)-f(t)}{\Delta t} = v$$

The term at the left side of the equation describes the change of a function that depends on a variable t when t also changes at a small amount Δt . What is $f(t)$? It is of course a way. Because speed is defined as way per time interval ($v = \text{distance} / \text{time}$) we can introduce this into our above equation and get

$$\frac{v(t+\Delta t)-vt}{\Delta t} = \frac{\Delta d}{\Delta t} = v = \text{const}$$

Note that in the latter equation the term $v(t+\Delta t)$ is a product, but in the function above $f(t+\Delta t)$ a function. The brackets enclose therefore once a term to be multiplied (for simplicity the multiplication sign was left out) and once denotes a list of variables of a function.

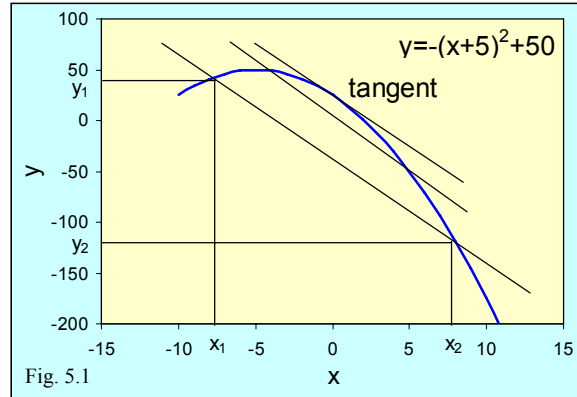
The latter equation may be generalized. If we have any function $y = f(x)$ we can look at the difference between two y -values $y_2 - y_1$ depending on the respective difference of the independent variable $x_2 - x_1$. This is graphically shown in the first Figure (Fig. 5.1). We may describe the change of a function $f(x)$ by the quotient $\Delta y / \Delta x$. This is the slope of the straight line through the two data points (x_1, y_1) and (x_2, y_2) .

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{\Delta y}{\Delta x} = \frac{f(x_1 + \Delta x) - f(x_1)}{\Delta x} \quad (5.1)$$

Often, changes of our function depend on the actual value of the function. Consider the population growth of a bacterium and assume that the bacterium divides at a constant rate r of very 20 min. In this case the total number of bacteria at time $t+\Delta t$ depends directly on the total number at time t . $N(t+\Delta t) = rN(t)$. The change in time is proportional to N :

$$\frac{N(t+\Delta t)-N(t)}{\Delta t} = rN(t)$$

To compute the population size of a bacterium at time $t+\Delta t$ we are therefore interested in the tangent at point $N(t)$. We can compute this if we take x_1 as t and x_2 as a value that is only very very very slightly larger



than x_1 . Our Δt should therefore be very small. In effect it should be nearly 0. It can't be equal to zero because a division through 0 is not defined but we may assume that Δt goes to 0. We define

$$\lim_{\Delta t \rightarrow 0} \Delta t = 0.$$

Now we can define a **derivative of a function** (that is the slope of the tangent through the point (x,y)) and therefore the change of the function at point (x,y) as follows

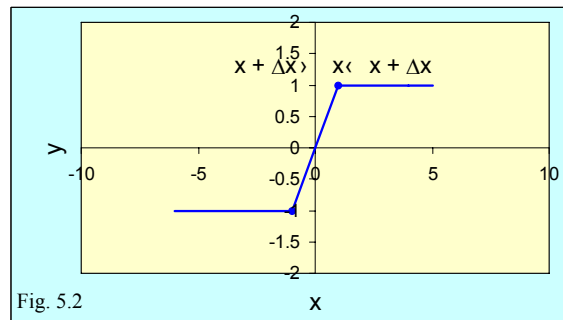
$$\lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{(x + \Delta x) - x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \frac{dy}{dx} \tag{5.2}$$

The latter quotient is often also written as

$$\frac{dy}{dx} = \frac{df(x)}{dx} = f'(x) = \dot{y} \tag{5.3}$$

If dy/dx (the slope of the function) is positive, the function increases, if dy/dx is negative the function decreases. For $dy/dx = 0$ at point a no change occurs at a.

Now look at the next Figure 5.2. It shows a modified Dirac function. This is a simple function with two thresholds. For computing the derivative it should make no difference whether I come from the right side and have a positive Δx or whether I come from the left side (with a negative Δx). But in the case of the Dirac function the direction matters. Below $x = -1$ the function has at every point the derivative 0 (the slope of the function is



0). The same holds for values above 1. But between -1 and $+1$ the slope has a constant value of 1, the derivative of the function $y = x$ is exactly 1. So at the points -1 and $+1$ we would be able to compute two derivatives with values 0 and 1. This makes no sense and we modify our above definition in such a way that we exclude that points where none or more than 1 derivative values would be possible. These are points where our function is not **continuous**. Look back to the tangent function. This function has an discontinuity at $\alpha = 90^\circ = \pi/2$. At these points the derivative of the tangent would not be defined.

We are now able to compute some simple derivatives. Consider the function $y = f(x) + g(x)$. The functions $f(x)$ and $g(x)$ are two functions that depend on x . We want to compute the derivative dy/dx .

$$\begin{aligned} \frac{dy}{dx} &= \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x) + g(x + \Delta x) - g(x)}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} + \lim_{\Delta x \rightarrow 0} \frac{g(x + \Delta x) - g(x)}{\Delta x} \end{aligned}$$

Therefore

$$\frac{d[f(x) + g(x)]}{dx} = \frac{dy}{dx} = \frac{df(x)}{dx} + \frac{dg(x)}{dx} \tag{5.4}$$

This is the **summation rule** for derivatives. In other words, the derivative of a sum is the sum of both derivatives if both depend on the same variable.

Consider next a product $y=f(x)*g(x)$. We get

$$\frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) * g(x + \Delta x) - f(x) * g(x)}{\Delta x}$$

Now, denote $f(x)$ as u and $g(x)$ as v . That gives

$$\begin{aligned} \frac{dy}{dx} &= \lim_{\Delta x \rightarrow 0} \frac{(u + \Delta u)(v + \Delta v) - uv}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{uv + v\Delta u + u\Delta v + \Delta u\Delta v - uv}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} v \frac{\Delta u}{\Delta x} + \lim_{\Delta x \rightarrow 0} u \frac{\Delta v}{\Delta x} + \lim_{\Delta x \rightarrow 0} \frac{\Delta u\Delta v}{\Delta x} \end{aligned}$$

$\lim_{\Delta x \rightarrow 0} \Delta u / \Delta x$ is du/dx , the derivative of u and for the function v holds the same. Additionally, it can be proven (and is intuitively obvious) that the last summand $\Delta u \Delta v / \Delta x$ goes against 0. From this we get

$$\frac{dy}{dx} = v \frac{du}{dx} + u \frac{dv}{dx} \quad (5.5)$$

Or, if we go back to our first notation

$$y = f(x)g(x) \quad \rightarrow \quad \frac{dy}{dx} = f(x) \frac{dg(x)}{dx} + g(x) \frac{df(x)}{dx} \quad (5.6)$$

This is a simple rule (called **product rule** of differentiation) that tells us how to compute the derivative of a product.

If we have a quotient of two functions $y = f(x) / g(x)$ we get with the same logic the **quotient rule** of differentiation

$$y = \frac{f(x)}{g(x)} \quad \rightarrow \quad \frac{dy}{dx} = \frac{g(x) \frac{df(x)}{dx} - f(x) \frac{dg(x)}{dx}}{g^2(x)} \quad (5.7)$$

If we set in 5.7 $f(x) = 1$ we get

$$\frac{d \frac{1}{g(x)}}{dx} = \frac{-1}{g^2(x)} \frac{dg(x)}{dx} \quad (5.8)$$

We have also to consider the case were two functions are nested. This is a function of the type $y = f[g(x)]$. To get the derivative we denote $g(x) = u$. Then we have dy/du and du/dx . For simplicity we multiply the right side of the derivative with du and get

$$\begin{aligned} y = f[g(x)] \quad \rightarrow \quad \frac{dy}{dx} &= \frac{dy}{du} * \frac{du}{dx} \\ \frac{dy}{dx} &= \frac{df(u)}{du} * \frac{dg(x)}{dx} \end{aligned} \quad (5.9)$$

In other words, the derivative of a nested function is the product of the derivatives of all included functions. This is the **chain rule** of differentiation. At last, we have to consider the **inverse of a function**.

A function $y = f(x)$ has the inverse $x = g(y)$. We can simply denote

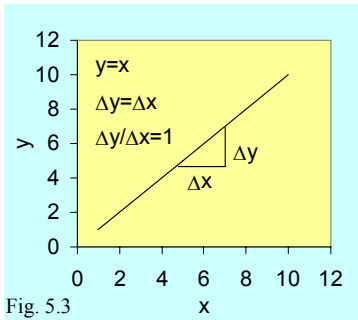


Fig. 5.3

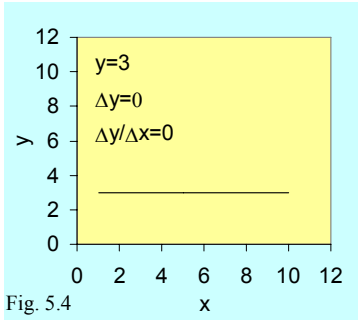


Fig. 5.4

$$\frac{dy}{dx} = \frac{df(x)}{dx} = \frac{1}{\frac{df(x)}{dy}} = \frac{1}{\frac{dg(y)}{dy}} = \frac{1}{dy}$$

(5.10)

and see that the derivative of an inverse function is the inverse derivative of the original function.

Now we have six simple rules to compute derivatives. With these rules it is possible to get derivatives from all functions that allow these derivatives at least at some points to be computed. Derivatives are again functions and we may derivate them too. Then we get derivatives of higher orders denoted as f'' , f''' or $f^2(x) / dx$, $f^3(x) / dx$ and so on.

Now we are ready to compute derivatives. Two simple cases can immediately be inferred from the respective plots. The function $y = mx$ is linear and has a constant slope m (Fig. 5.3). $dy / dx = m$. For $y = x$ holds that $dy / dx =$

1. Fig. 5.4 instead shows that the derivative of a constant is zero. A constant has no change on the y-axis.

Let's consider other important examples. We know already the number e , Euler's number, as the base of the natural logarithm. e has the value 2.718282... Why is e so important and why has this number such a complicated value? Because of a very nice feature. e is chosen in that way that

$$\frac{de^x}{dx} = e^x$$

(5.11)

In other words the function e^x gives in every point for that this function is defined its own derivative (Fig. 5.5). This makes many things very easy. With this definition we can compute other derivatives too. What is $d \ln(x) / dx$? Our rules about inverse functions tell us that $dy/dx = 1/(dx/dy)$. Therefore

$$\frac{d \ln(x)}{dx} = \frac{1}{\frac{de^y}{dy}} = \frac{1}{e^y} = \frac{1}{x}$$

(5.12)

Our rule gave us immediately the derivative of a logarithm (Fig. 5.6).

We also get

$$y = ax^b$$

$$\frac{dy}{dx} = \frac{d(e^{\ln a + b \ln x})}{dx} = \frac{b}{x} ax^b = abx^{(b-1)}$$

(5.13)

and

$$y = ab^x$$

$$\frac{dy}{dx} = \frac{d(e^{\ln a + x \ln b})}{dx} = ab^x \ln b$$

(5.14)

Additionally, eq. 5.12 provides us with another important rule for differentiation.

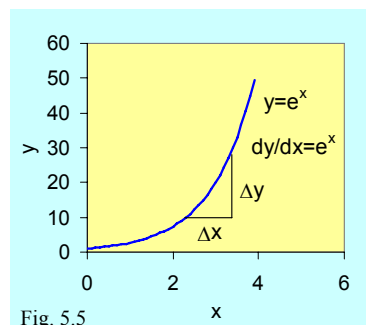


Fig. 5.5

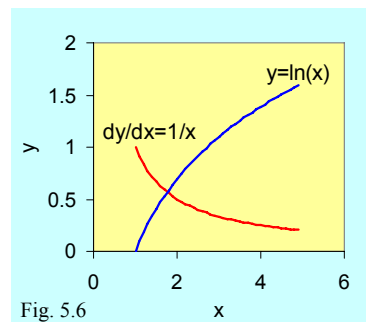


Fig. 5.6

Suppose we need $y' = d[u(x)v(x)] / dx$. The chain rule tells us that

$$\frac{d \ln(y)}{dx} = \frac{1}{y} \frac{dy}{dx} \rightarrow \frac{1}{y} \frac{dy}{dx} = \frac{1}{u} \frac{du}{dx} + \frac{1}{v} \frac{dv}{dx} \rightarrow \frac{dy}{dx} = uv \left(\frac{1}{u} \frac{du}{dx} + \frac{1}{v} \frac{dv}{dx} \right)$$

Of course today differentiation is often done by math programs. There are also some internet pages that offer calculus. The next Figure shows some simple examples of functions that would surly take some time to be evaluated by hand. With *Mathematica* it takes a few seconds.

Differential calculus runs into difficulties if we want to differentiate a function that is not defined or is zero or becomes infinite at a certain point . Consider the important function

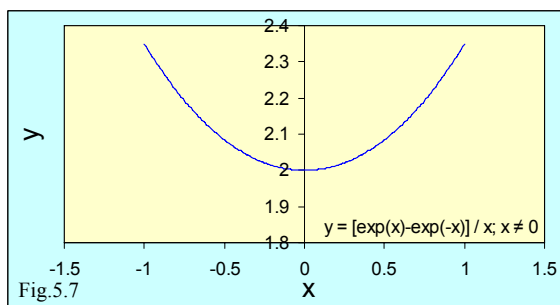
$$y = \frac{e^x - e^{-x}}{x}$$

This function is not defined for $x = 0$. We would get $0 / 0$ and we would be able to compute two derivative values depending on the side from which we



come. But the Figure 5.7 (a plot of the function) indicates that the function should have a lower boundary at this point. We are frequently interested in such boundaries. From the plot we expect a value of 2. Can we compute the boundary of the function at this point? For such tasks the **rule of Guillaume de l'Hospital** (a French mathematician who wrote the first textbook on calculus, 1661-1704) helps. It states that if two tions $f(x)$ and $g(x)$ vanish at a certain point a the following equation holds:

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)} \tag{5.14}$$



To show this we need $f(a) = g(a) = 0$

Therefore

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{h \rightarrow 0} \frac{f(a+h)}{g(a+h)} = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{g(a+h) - f(a)} = \lim_{h \rightarrow 0} \frac{\frac{f(a+h) - f(a)}{h}}{\frac{g(a+h) - f(a)}{h}} = \frac{f'(a)}{g'(a)}$$

For the above example we get

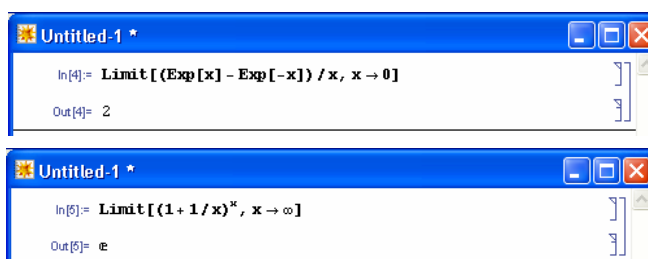
$$\lim_{x \rightarrow 0} \frac{f'(x)}{g'(x)} = \lim_{x \rightarrow 0} \frac{e^x + e^{-x}}{1} = \frac{1+1}{1} = 2$$

The rule of de l'Hospital gives the expected value of the function at $x = 0$. Of course, math programs are also able to compute limits.

Another example of this important rule.

What is

$$\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x} \right)^x$$



This is not a quotient necessary to apply the rule of de l' Hospital. So, we have to transform the function into an appropriate form. We take

$$A = \lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x \rightarrow A = e^{\lim_{x \rightarrow \infty} x \ln\left(1 + \frac{1}{x}\right)}$$

Now, we have to transform the exponent in that way that it matches the requirements of the above rule. Functions that go to infinity have to be transformed by taking $1 / f(x)$. We transform

$$\ln(A) = \lim_{x \rightarrow \infty} x \ln\left(1 + \frac{1}{x}\right) = \lim_{x \rightarrow \infty} \frac{\ln\left(1 + \frac{1}{x}\right)}{\frac{1}{x}} = \lim_{y \rightarrow 0} \frac{\ln(1+y)}{y}; \text{ for } y = 1/x$$

We differentiate and get

$$\lim_{y \rightarrow 0} \frac{f'(y)}{g'(y)} = \lim_{y \rightarrow 0} \frac{1/1}{1} = 1$$

Therefore

$$A = \lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x \rightarrow A = e^{\lim_{x \rightarrow \infty} x \ln\left(1 + \frac{1}{x}\right)} = e^1 = e = \lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x \quad (5.16)$$

This is another remarkable feature of the number e.

Now consider $y=(e^n-1)/n$. We apply the **rule of de l'Hospital** and get

$$\lim_{n \rightarrow 0} \frac{e^n - 1}{n} = \lim_{n \rightarrow 0} \frac{e^0}{1} = 1 \quad (5.17)$$

We get the next important limit again using the de l'Hospital rule.

$$\lim_{n \rightarrow 0} \frac{\ln(1+n)}{n} = \lim_{n \rightarrow 0} \frac{1/(1+1)}{1} = 1 \quad (5.18)$$

The above examples all dealt with **explicit functions**, functions that could be written in the explicit form $y = f(x)$. Unfortunately, this is quite often not possible and we have to use functions in an **implicit form** where it is not possible to solve for y. For instance the function

$$\sqrt{x^2 + y^2} = \ln(x + y)$$

cannot be solved for y. Is it possible do compute the derivative dy / dx ? It is and we have to use the chain rule to reach in an **implicit differentiation**. We take

$$\frac{d}{dx} \left(\sqrt{x^2 + y^2} - \ln(x + y) \right) = \frac{d(0)}{dx} = 0$$

The chain rule gives us

$$\frac{4x + 4y \frac{dy}{dx}}{\sqrt{x^2 + y^2}} - \frac{\frac{dy}{dx}}{x + y} = 0$$

and from this we compute

$$4x(x+y) + 4y(x+y) \frac{dy}{dx} - \frac{dy}{dx} \sqrt{x^2 + y^2} = 0$$

and

$$\frac{dy}{dx} = \frac{4x(x+y)}{\sqrt{x^2 + y^2} - 4y(x+y)}$$

This was of course a quite complicated example but it trained us a little bit how to use the chain rule.

Let's try an easy example. Consider the implicit function $xy = 1$? We get

$$\frac{d}{dx}(xy) = x \frac{dy}{dx} + y \frac{dx}{dx} = \frac{d(1)}{dx} = 0 \rightarrow \frac{dy}{dx} = \frac{-y}{x}$$

In this case we can check our result. $y = 1 / x$ and we infer $dy / dx = -1 / x^2$. The quotient rule gives the same result.

Now we have to deal with **minima** and **maxima** and **points of inflection**. Fig. 5.8 shows a simple algebraic function. We need the points $y = f(x)$ where a local maximum and minimum occurs.

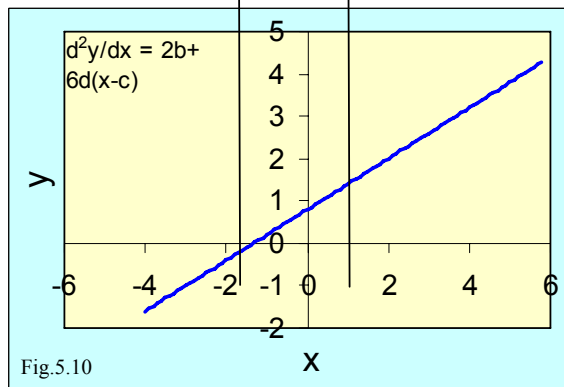
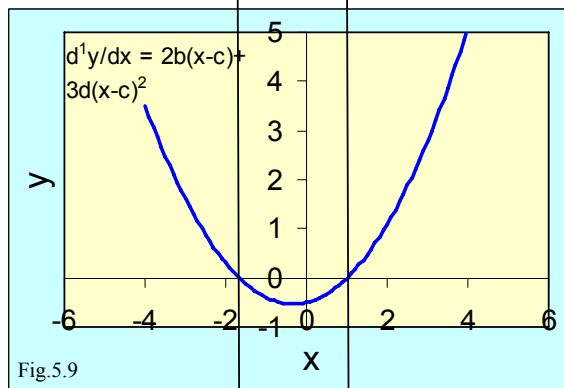
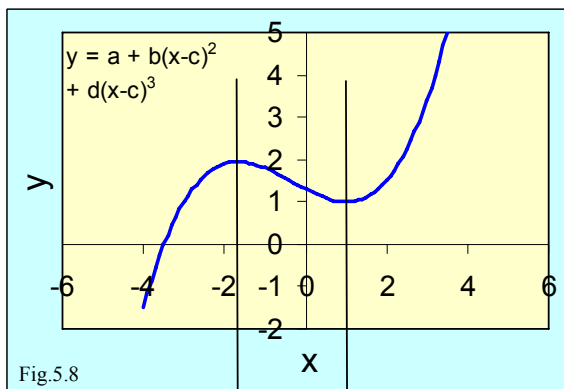


Fig. 5.9 immediately shows that at these points the first derivative is exactly zero. This is clear because at these two points no change occurs. Hence by setting the first derivative of a function to zero we can compute local maxima and minima.

How to differentiate between a maximum and a minimum? Fig. 5.10 shows that a maximum occurs if the second derivative is negative. For a point to be a minimum the second derivative has to be larger than zero. This is intuitively logical because the derivative gives the slope of a function at all points x . If this slope is zero, the derivative should also be 0. Setting the derivative to zero is therefore a convenient method to obtain minimum and maximum values of a function. On the other hand, we saw that if we already deal with a derivative, that means with a function that describes changes of a variable we have to set the function to zero to obtain that points where no further changes occur. We call them **stationary points**. These are often points of **equilibrium**.

From Figs. 5.8. to 5.10 we got a method to determine maxima and minima of functions. Compute the first and second derivative (if they exist) and look where the first derivative is zero and whether the second derivative is at this point below or above 0. Of course, this is not a mathematical proof in a strict sense. It is more a sketch

of a proof where each step has to be worked out in detail. However, our aim is not to make strict mathematics, our aim is only to understand basic relations and for this task simplified ‘proofs’ are much better suited.

Now look at Fig. 5.11. We consider the function

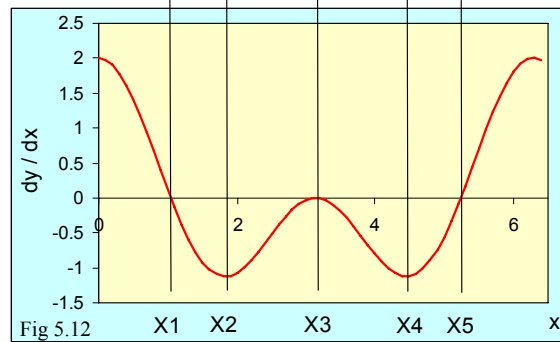
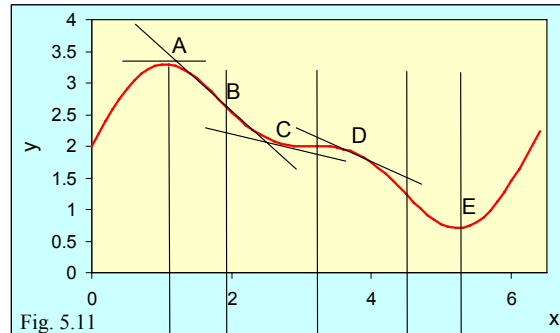
$$y = \sin(x)(1 + \cos(x)) + 2$$

The plot reveals a local maximum at point A and a local minimum at point E. The function decreases at B. Hence $y'(B) < y'(A) = 0$. At C however y still decreases but more slowly. Hence $y'(C) > y'(B)$. In total $y'(A) > y'(B) < y'(C)$. At D the function again decreases more than at C. Therefore between A and C must be a point where the derivative y' becomes maximal. This is the point where the function changes its direction. At these points the function decreases or increases maximal. For biological processes these points for instance might denote changes in temporal patterns. We find these points by studying the first and second derivative. At inflection points the first derivative has a local maximum or minimum, the second derivative becomes therefore zero. (Fig. 5.12). In our case we get

$$y' = \cos(x)(1 + \cos(x)) - \sin^2(x) \quad \text{and}$$

$$y'' = -4 \cos(x) \sin(x) - \sin(x)$$

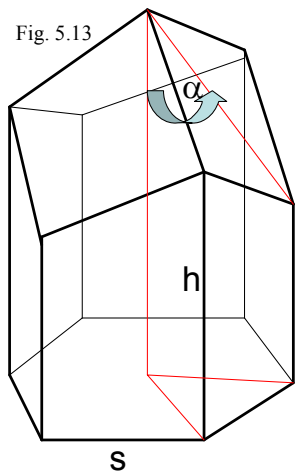
This function is zero at $\cos(x) = -1/4$. By setting $y'' = 0$ *Mathematica* finds the first four solutions and solves arc cos (-1/4) numerically. Other points of inflections would be at $n\pi + \arccos(-1/4)$. Fig. 5.12 shows local maxima and minima of y' at these points.



```

Untitled-1 *
In[2]:= D[Sin[x] (1 + Cos[x]) + 2]
Out[2]= Cos[x] (1 + Cos[x]) - Sin[x]^2
In[3]:= D[Cos[x] (1 + Cos[x]) - Sin[x]^2]
Out[3]= -3 Cos[x] Sin[x] - (1 + Cos[x]) Sin[x]
In[4]:= Solve[-3 Cos[x] Sin[x] - (1 + Cos[x]) Sin[x] == 0, x]
Solve::ifun :
Inverse functions are being used by Solve, so some
solutions may not be found; use Reduce
for complete solution information. More_
Out[4]= {{x -> 0}, {x -> -ArcCos[-1/4]}, {x -> ArcCos[-1/4]}}
In[6]:= N[ArcCos[-1/4]]
Out[6]= 1.82348
    
```

Fig. 5.13



A simple, but well known example, how to solve the minimum problem in biology (I took this example from Portenier, Gromes, *Mathematik für Biologen und Humanbiologen*, Marburg 2003). Bee combs have a hexagonal shape. Is this shape optimal? We can reformulate this problem by asking at what shape honey combs take the least amount of wax. Hence the prob-

```

Untitled-1 *
In[12]:= D[Sqrt[3 - Cos[x]] / Sin[x]]
Out[12]= 1 - (Sqrt[3 - Cos[x]] Cot[x] Csc[x])
In[13]:= Simplify[1 - (Sqrt[3 - Cos[x]] Cot[x] Csc[x])]
Out[13]= Csc[x] (-Sqrt[3] Cot[x] + Csc[x])
In[14]:= Solve[Csc[x] (-Sqrt[3] Cot[x] + Csc[x]) == 0, x]
Solve::ifun :
Inverse functions are being used by Solve,
so some solutions may not be found.
Out[14]= {{x -> -ArcCot[1/Sqrt[2]]}, {x -> ArcCot[1/Sqrt[2]}}
In[16]:= N[ArcCot[1/Sqrt[2]]]
Out[16]= 0.955317
    
```

lem is to find a hexagonal comb shape of which the surface is minimal. Because of the symmetrical shape, the surface can be described by one angle α (Fig. 5.13). The surface of a comb is given by

$$S(\alpha) = 6hs + \frac{3s^2}{2} \frac{\sqrt{3} - \cos(\alpha)}{\sin(\alpha)}$$

Hence, to find the maximum, the first derivate has to be zero. We apply a math program and solve for α (instead of α we use x). First, we compute the derivative of the last term. Then we solve for x. At last, we compute the numerical solution. This is 0.955. These are radians, the standard output of math programs. Hence $0.955 / 2 \pi = \alpha / 360^\circ$; $\alpha = 54.7^\circ$. This is the angle at which the amount of wax would be minimal. Indeed, real bee caves have an angle that is very similar to this solution.

6. Summing up

Up to now we looked at functions that could be written in precise equations. However, this is often not possible. We have instead to deal with a series of single equations that together model a whole process. Look at the following simplified example (Fig. 6.1). Assume in a certain habitat occur S_0 species of animals in one year t_0 . At time t_1 a number of species immigrated from the so-called general **regional species pool** P and some species emigrated or went locally extinct. Assume further the annual local species turnover p (the proportion of new species) is the sum of emigration and immigration rates and is proportional to the mean number of annual species. The latter is of course S_0 .

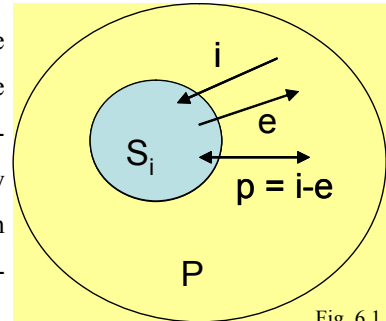


Fig. 6.1

$$S_{immigration} \propto S_i = iS_0$$

$$S_{emigration} \propto S_i = -eS_0$$

Now we are interested in the total number of different species in our habitat. If we have S_0 species at time 0 the number of species after a certain time step (for instance one generation) is

$$S_1 = S_0 + iS_0$$

Note that we only need the immigration rate I because only immigrations provide new species. Additionally, p depends on the regional species pool. If S_0 would be large in relation to P the number of new species able to immigrate would be small. This can be modeled by a multiplicative term

$$iS_0 \propto \frac{P - S_i}{P}$$

From these assumptions we can model the cumulative number of species S_1 ($= S_0 +$ the new species) by the following equation

$$S_1 = S_0 + iS_0 \frac{P - S_0}{P}$$

For the next generation we get

$$S_2 = S_1 + iS_0 \frac{P - S_1}{P}$$

For time n we have therefore

$$S_n = S_{n-1} + iS_0 \frac{P - S_{n-1}}{P}$$

This is a so-called **recursive equation** because the outcome for step n depends on the outcome for step $n-1$ and it is not immediately possible to reduce this process so that the outcome at step n can be derived only from the initial settings. Such processes are very common in biology and we have to look a little bit closer to them. Our process is a so-called first order recursive process because in the term of the right side of the equation occurs only S_{n-1} . If also S_{n-2} would occur we would speak of a second order recursive process and so on. Additionally, there are situations where recursive equations are **nested**. In these cases the elements $E(n-k)$ are itself functions of previous elements $E(n-m)$ ($m > k$).

Our recursive process is difficult to solve, but if the local species numbers S_i would be small in relation to the regional species pool the quotient $(P-S_i)/P$ would be nearly 1 and the last equation would reduce to

$$S_n \approx S_{n-1} + iS_0 \approx S_{n-2} + iS_0 + iS_0 \dots \approx S_0 + niS_0 \approx S_0(1 + ni) \tag{6.1}$$

Unfortunately, most often recursive processes do not allow a reduction to be made and we have to solve them numerically. This is not difficult with modern spreadsheet programs. The Table beside shows, for instance, a solution of the above process using an Excel spreadsheet.

	A	B	C	D	E
1	Equilibrium species number S_0			100	
2	Regional species pool			1000	
3	Immigration rate i			0.1	
4	Time	0		=+D1	
5		=+B4+1	:+D4+\$D\$3*\$D\$1*((\$D\$2-D4)/\$D\$2		
6		=+B5+1	:+D5+\$D\$3*\$D\$1*((\$D\$2-D5)/\$D\$2		
7		=+B6+1	:+D6+\$D\$3*\$D\$1*((\$D\$2-D6)/\$D\$2		
8		=+B7+1	:+D7+\$D\$3*\$D\$1*((\$D\$2-D7)/\$D\$2		

Next we look at mathematical series. Consider, for instance the following, so-called **geometric series**

$$N = a + ax + ax^2 + ax^3 + ax^4 \dots \sum_{i=1}^n ax^i \tag{6.2}$$

Is it possible to infer the sum? We multiply both sides with x and get

$$xN = x \sum_{i=0}^n ax^i = (ax + ax^2 + \dots ax^n) + ax^{n+1} = \sum_{i=0}^n ax^i - a + ax^{n+1} = N - a + ax^{n+1}$$

Fortunately, our series reduces to an equation that can easily be solved

$$N = \sum_{i=0}^n ax^i = \frac{a - ax^{n+1}}{1 - x} \tag{6.3}$$

Of course, our function is not defined for $x = 1$. If $x < -1$ or $x > 1$ the above series goes to infinity if n goes to infinity. But what is if $-1 < x < 1$? If n goes to infinity, the term x^n becomes smaller and smaller. Therefore the summands become also smaller and smaller as n raises. Is there a limit?



Look

$$N = \lim_{n \rightarrow \infty} \sum_{i=1}^n ax^i = \frac{a}{1-x} (1 - x^{n+1})$$

If n goes to infinity the term x^{n+1} goes to 0 and $(1-x^{n+1})$ goes to 1. Therefore

$$N = \frac{a}{1-x} \tag{6.4}$$

We see that our infinite series has a finite sum. *Mathematica* gives the same result.

Many such series have finite sums and sometimes it is important to determine whether there is such a sum or not. There are many criteria for this which involve the partial sums of the series We denote

$$S_n = \sum_{i=1}^n a_i$$

with a_i being the elements of the series. Our aim is now to look whether all elements S_n are below a certain

value a . This is equivalent with the **convergence** of the whole series. Otherwise the series **diverges**. The most important rule is the following: If a series contains only positive summands a the series converges to a finite sum if there is a positive number $q < 1$ for what all quotients $S_{n+1} / S_n < q$. This is the co-called **rule of d'Alembert** (a French mathematician, 1717-1783). For instance the quotient S_{n+1} / S_n of the series $1+1/2+1/3+1/4...$

$$\lim_{n \rightarrow \infty} \frac{\sum 1/(n+1)}{\sum 1/n} \rightarrow 1$$

goes for large n to 1. Hence the series does not converge, it diverges.

The next important criterion is that of **Cauchy**. A series a is convergent if for all n and $i, j > n$ an ε exists for which

$$|S_j - S_i| < \varepsilon \tag{6.5}$$

Applied again to the harmonic series gives

$$|S_{2n} - S_n| = \sum_{i=1}^{2n} \frac{1}{i} - \sum_{i=1}^n \frac{1}{i} = \sum_{i=n+1}^{2n} \frac{1}{i} > n \frac{1}{2n} > \frac{1}{2}$$

Because this holds for all n the harmonic series diverges. However the series $1+1/4+1/9+1/16...$ converges

$$|S_{2n} - S_n| = \sum_{i=1}^{2n} \frac{1}{i^2} - \sum_{i=1}^n \frac{1}{i^2} = \sum_{i=n+1}^{2n} \frac{1}{i^2} > n \frac{1}{n^2} < \frac{1}{n}$$

Another criterion of Cauchy is the following. If

$$\lim_{n \rightarrow \infty} |a_n|^{1/n} < 1 \tag{6.6}$$

the respective series converges. Again, we test this rule with an example. We take the series

$$\sum_{i=1}^{\infty} \left(\frac{n-1}{n} \right)^{n^2}$$

Therefore, we determine whether

$$\lim_{n \rightarrow \infty} \left(\frac{n-1}{n} \right)^n < 1$$

We get

$$\lim_{n \rightarrow \infty} \left(\frac{n-1}{n} \right)^n = \lim_{n \rightarrow \infty} \left(\frac{n}{n+1} \right)^n = \lim_{n \rightarrow \infty} \left(\frac{1}{1+1/n} \right)^n = \frac{1}{e} < 1$$

Therefore, our series converges

A last criterion is that of **Leibniz**. A series of elements with alternating signs converges if the absolute values steadily converge to zero. For instance $1-x+x^2-x^3...$ gives

$$\lim_{n \rightarrow \infty} x^n = 0; 0 < x < 1$$

The series converges for $0 < x < 1$.

Our criteria don't tell us the sum of our series. It is often very difficult to find the sum of an infinite

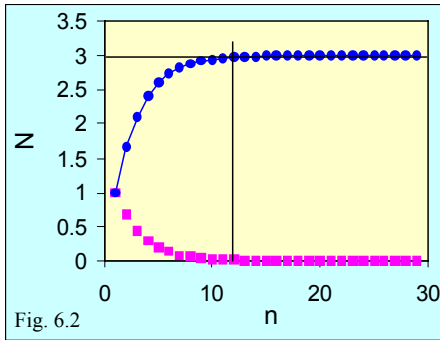


Fig. 6.2

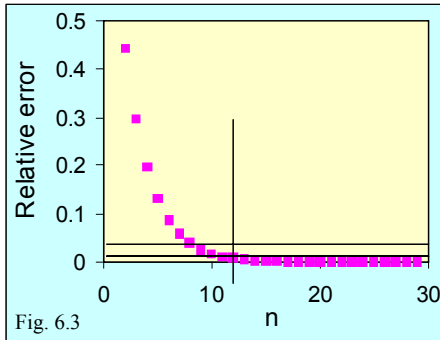


Fig. 6.3

series. In these cases we have to approximate the sum by a computer simulation.

Sometimes it is of interest how fast a series **converges**. For instance, in statistics this problem is closely related to the problem how many probes we have to take or experiments to make (in general how large the **sample size** has to be). This is a problem of the so-called **power analysis** with which we will deal in the statistics lecture. Look at Fig. 6.2. It shows the elements of $y = 1 / (1.5)^n$ and the sum of the elements. Our above equation 6.3 predicts for this series a sum of $1 / (1-2/3) = 3$. We see that our series very quickly converges to this limit. It goes asymptotically to 3. Such **asymptotes** are very often found in nature. Fig. 6.3 shows the relative error we make when we deal only with a limited number of summands as an approximation of the real value. This relative error is the following quotient

$$\text{Error} = \frac{|\text{real value} - \text{approximation}|}{\text{real value}} = 1 - \frac{\text{approximation}}{\text{real value}} \tag{6.7}$$

In our example we need only 12 summands for our relative error to fall below 1%. But even if we would take only 8 summands, our relative error would already be below 5%. We say that the series converges fast.

Another example. One of the most often used series in statistics (but also otherwise) is the **Newton series** $(a + b)^n$. For $n = 2$ it is easy to solve $(a + b)(a + b) = a^2 + 2ab + b^2$. Let's add another factor $(a + b)(a + b)(a + b) = (a^2 + 2ab + b^2)(a + b) = a^3 + 3a^2b + 3ab^2 + b^3$. If we try out a large number of n we will find a nice regularity. We see that the exponents of b continuously rise from 0 to n , that of a continuously decrease from n to 0. The constants follow a regularity called the **Pascal triangle** (after the French mathematician and philosopher Blaise Pascal, 1623-1662, although this triangle already appeared in a Chinese textbook from about 1100 b. c.) that is shown in Fig. 6.4. The elements of each

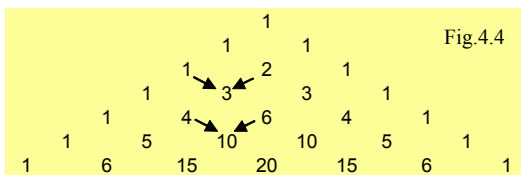


Fig.4.4

row are the sums of the elements above. In the statistics lecture we will see that we also can compute these constants from the number of combinations of elements. In general the above function is the sum of the series

$$(a + b)^n = \sum_{i=0}^n \frac{n!}{i!(n-i)!} a^i b^{n-i} \tag{6.8}$$

$n!$ is again the product of all i from $i = 1$ to n . This is the so-called **binomial function** and later we will deal with it in detail. How fast this series converges against $(a + b)^n$ depends of course on the values of a , b and n . For a special case where $a = 1$ and x is well below 1 the series converges very fast and we can reduce it to the first elements. The Figure above (Fig. 6.5)

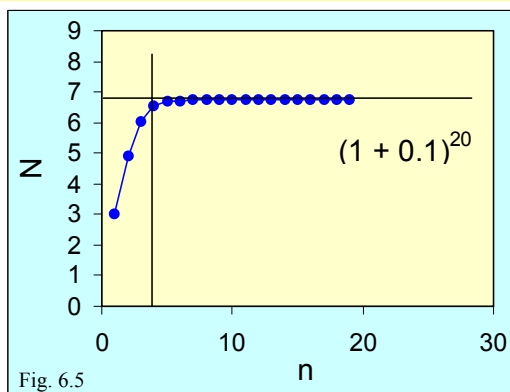


Fig. 6.5

shows that in the case of $x = 0.1$ the first four elements are sufficient for the relative error to become less than 1%.

The last examples tried to develop a given function into a series of functions of which it is the sum.

$$f(x) = f_1(x) + f_2(x) + f_3(x) \dots \sum_{n=0}^N f_n(x)$$

At first sight this seems not to be a clever and easy method to handle functions. But in reality it is of great importance. Often, it is impossible to deal with complicated functions, for instance, to obtain their roots or to integrate them. In such cases we may try to approximate the function by a series.

We deal with two important general methods that allow a large number of functions to be developed into series. Firstly, we assume that we deal only with such functions that can be developed into an **algebraic function** of the form

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 \dots = \sum_{i=0}^{\infty} a_i x^i \quad (6.9)$$

This is a very general form of an algebraic function. Our task is now to determine the constants a_i . This algebraic function and therefore $f(x)$ is also assumed to have derivatives of all orders at point 0. For $x = 0$ follows of course that $f(0) = a_0$, the first constant. Consider now the first derivative of $f(x) = f'(x)$. This is

$$f'(x) = 0 + a_1 + 2a_2x + 3a_3x^2 \dots + ia_i x^{i-1} \dots = \sum_{i=0}^{\infty} ia_i x^{i-1}$$

Now, we consider only the point $x=0$. This results immediately in $a_1 = f'(0)$. a_2 is of course $f''(0) / 2$, $a_3 = f'''(0) / (3*2)$, $a_4 = f^{(4)}(0) / (4*3*2)$, and $a_n = f^{(n)}(0) / n!$.

We now introduce these values of a_n into the original algebraic function and get

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 \dots = \sum_{i=0}^{\infty} \frac{f^{(i)}(0)}{i!}x^i \quad (6.10)$$

Note that in this latter series $0!$ occurs. By definition is $0! = 1$. $f^{(0)}(0) = f(0)$.

Now, we know a first very important way to develop any appropriate function (a function for which the above assumptions hold) into an infinite series. This series is called **Mac Laurin series**.

Let's look what this series gives for certain important functions. Consider our already well known function $y = e^x$. Remember that $d(e^x) / dx = e^x$.

$$e^x = e^0 + \frac{e^0}{1!}x + \frac{e^0}{2!}x^2 \dots = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} \dots = \sum_{i=0}^{\infty} \frac{x^i}{i!} \quad (6.11)$$

For $x = 1$ we get

$$e = e^0 + \frac{e^0}{1!} + \frac{e^0}{2!} \dots = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} \dots = \sum_{i=0}^{\infty} \frac{1}{i!} \quad (6.12)$$

We can generalize our result for other values of x . This is done by a so-called **Taylor series** (after the British mathematician Brook Taylor, 1685-1731; although others like Leibniz and Newton used similar series prior). Assume that we generalize the way our function may be developed into an algebraic function

$$f(x) = a_0 + a_1(x-b) + a_2(x-b)^2 + a_3(x-b)^3 \dots = \sum_{i=0}^{\infty} a^i (x-b)^i$$

Now, we apply the same logic as above and use $f^n(x-b)$. We do this by setting $x = b$ and get immediately

$$f(x) = f(b) + \frac{f^1(b)}{1!}(x-b) + \frac{f^2(b)}{2!}(x-b)^2 + \frac{f^3(b)}{3!}(x-b)^3 \dots = \sum_{i=0}^{\infty} \frac{f^i(b)}{i!}(x-b)^i \quad (6.13)$$

Let's look at two examples how to use Taylor and Mac Laurin series. What about $y = (a+x)^n$, the Newtonian or binomial equation? We take a Mac Laurin series and get

$$(a+x)^n = f(0) + \frac{f^1(0)}{1!}x + \frac{f^2(0)}{2!}x^2 \dots = a^n + \frac{n(a)^{n-1}x}{1!} + \frac{n(n-1)(a)^{n-2}x^2}{2!} \dots = \sum_{i=0}^n \binom{n}{i} a^{n-i} x^i$$

The above function uses an important shortage that is defined in the following way

$$\binom{n}{i} = \frac{n!}{i!(n-i)!} \quad (6.14)$$

Our Mac Laurin expansion gives us an elegant (sketch of a) prove of the binomial function. One special case of this is very interesting. Consider $y = (1+x)^{-1}$. The above equation gives immediately

$$\frac{1}{(1+x)} = 1 - x + x^2 - x^3 + x^4 \dots = \sum_{i=0}^{\infty} (-1)^i x^i \quad (6.15)$$

An interesting special case occurs when a and x are both $1/2$. Then we get

$$\left(\frac{1}{2} + \frac{1}{2}\right)^n = \sum_{i=0}^n \binom{n}{i} \left(\frac{1}{2}\right)^n = 1$$

For x and $a = 1$ we get

$$(1+1)^n = \sum_{i=0}^n \binom{n}{i} (1)^n = 2^n$$

Taylor series have an immense importance in calculus because they are the basis for many approximations of complicated terms.

At the end we come back to recursive processes. We had already first order recursive processes. An important class of second order recursive functions are **Fibonacci series** (after Leonardo de Pisa known as Fibonacci, 1175? - 1240?, his *Liber Abaci* introduced the Indic decimal system and the Arabian numbers in Europe). A Fibonacci order has the following general structure

$$f(x) = f(x-1) + f(x-2) \quad (6.16)$$

Applied to natural numbers this gives

$$1=1+0$$

$$2=1+1$$

$$3=2+1$$

$$5=3+2$$



8=5+3
13=8+5
...

Eq. 6.16 can be rewritten in differential notation. For a sufficiently small ϵ we get

$$\frac{f(x) - f(x - \epsilon)}{x - (x - \epsilon)} = \frac{df(x)}{dx} = f(x - 2\epsilon) \approx f(x)$$

Fig. 6.6, Photo by US Agricultural plant database

This is a so-called differential equation (parts 9 and 10). The rate of change is proportional to its own value. Be-

cause $f(e^x) = e^x$ we get a solution

$$f(x) = ce^{ax} = c(e^a)^x = c\lambda^x$$

Introducing into 6.16 we have for $x = 2$

$$\lambda^2 = \lambda + 1 \rightarrow \lambda_{1,2} = \frac{1 \pm \sqrt{5}}{2}$$

Hence for large x the ratio $f(x)/f(x-1)$ converges to

$$\frac{f(x+1)}{f(x)} = \frac{c\lambda^{x+1}}{c\lambda^x} = \lambda = \frac{1 + \sqrt{5}}{2}$$

The ratio is well known as the **golden mean** or golden number.

There are some nice examples of (approximate) Fibonacci orders in nature. Best known is surely the geometric ordering of leafs on many plant stalks or the rosette ordering of crown leafs in flowers stands (Fig. 6.6). The reason for this is that the leafs are arranged in such a way to get a maximum amount of light. This is shown in the Figure 6.7.

The side length of areas arranged as a spiral follow the Fibonacci series. Hence ammonites and nautilus following logarithmic spirals also follow Fibonacci series as shown in Fig. 6.8 for a nautilus. There are many other optimization problems in nature that can be described by generalized Fibonacci orders. A nice collection of them contains R. Knotts Fibonacci site at <http://www.mcs.surrey.ac.uk/Personal/R.Knott/Fibonacci/fibnat.html>.

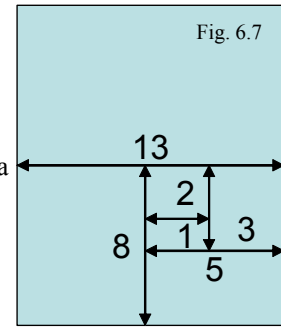


Fig. 6.7

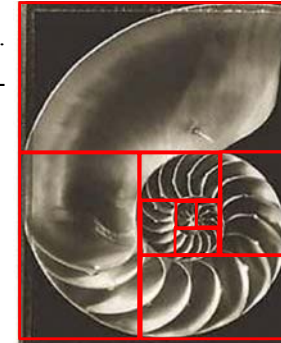
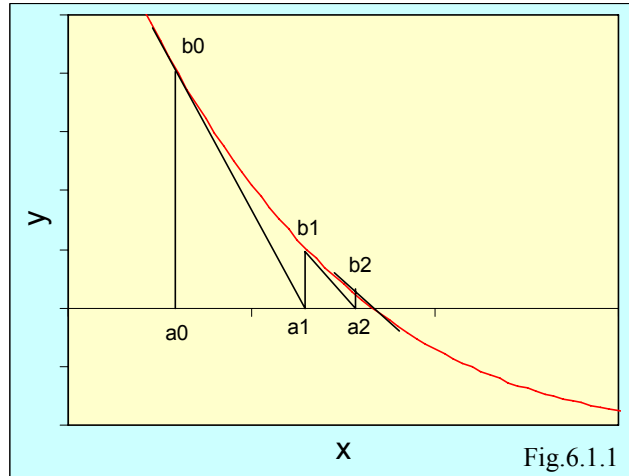


Fig. 6.8, Photo by T. Baril

6. 1 The Newton approximation

Often it is not possible to solve a function for its roots: $f(x) = 0$. The task can be solved numerically by the Newton approximation. Assume we have an initial guess about $x_0 = a_0$. At a_0 $f(x)$ has a value of b_0 . The derivative of $f(x)$ at $x = a_0$ has the new root a_1 and from $f(a_1)$ we get a new b_1 . After four steps the example of Fig. 6.1.1 gives an intercept of the x-axis very close to the root of $f(x)$. $f(a_1) = 0$. We get



$$\frac{f(a_1) - f(a_0)}{a_1 - a_0} = f'(a_0) \rightarrow a_1 = a_0 - \frac{f(a_0)}{f'(a_0)}$$

The same holds for all other points and we get the general Newton approximation

$$a_n = a_{n-1} - \frac{f(a_{n-1})}{f'(a_{n-1})} \tag{6.1.1}$$

For instance find the root of $y = e^{-2x} - x^2$. We have

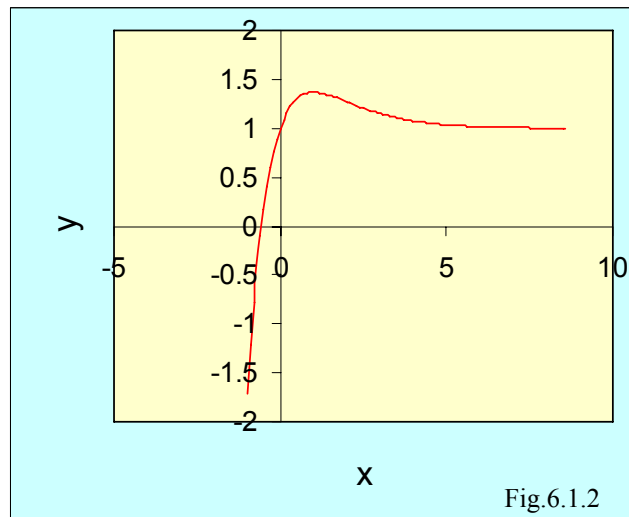
$$f'(e^{-2x} - x^2) = -2e^{-2x} - 2x$$

With $a_0 = 0.1$ and $b_0 = 0.81$ we get the following sequence $a_1 = 0.54$, $b_1 = 0.05$, $a_2 = 0.5673$, $b_2 = -0.00025$, and $a_3 = 0.5671$, $b_3 = 0$. We needed only three steps to obtain a nearly perfect result.

The Newton approximation heavily depends on the starting point. For instance we have to solve the equation $xe^{-x} = -1$ (Fig. 6.1.2). We use two starting points, $a_0 = 0.5$ and $a_0 = 1.5$. Hence

$$a_n = a_{n-1} - \frac{a_{n-1}e^{-a_{n-1}} + 1}{e^{-a_{n-1}} - a_{n-1}e^{-a_{n-1}}}$$

For $a_0 = 0.5$ we get the series $a_1 = -3.80$, $a_5 = -1.07$, $a_{10} = -0.567$, For $a_0 = 1.5$ we get the series $a_1 = 13.50$, $a_2 = 56435$, $a_3 = n.d.$ The crucial point is that a starting point above 1 is behind the local maximum at $x = 1$. Therefore for the Newton method to succeed we need starting points within the region defined by the nearest local maximum or minimum (if these occur).



7. The sums of infinities

We know already how to differentiate functions. Every mathematical operation has its counter operation to reverse it. The counter operation of addition is subtraction, of multiplication is division of derivation is what? We don't know. The simplest way to change this is to define such an operation. We define the **antiderivative** or the **indefinite integral** of a derivative $f(x)$ as

$$F(x) = \int f(x)dx \quad (7.1)$$

It is the operation that reverses the derivation. From this definition we get immediately some indefinite integrals

$$\begin{aligned} \int dx &= x + c \\ \int x^n dx &= \frac{x^{n+1}}{n+1} + c \\ \int e^{ax} dx &= \frac{e^{ax}}{a} + c \end{aligned} \quad (7.2)$$

Especially is true

$$\begin{aligned} \int af(x)dx &= a \int f(x)dx \\ \int f(x) + g(x)dx &= \int f(x)dx + \int g(x)dx \end{aligned} \quad (7.3)$$

To proof these equations you simply have to derivate them. But what is c in the above three equations? c is a constant, the so-called **integration constant**. We have to add this constant because derivating the right sides of the equations gives always the left side irrespective of the value of c . The integration gives therefore not an unequivocal result.

Derivating a function was a simple task. The integration is not so simple. There is not a fixed set of rules with which you can integrate all functions and you need a lot of experience and skill to find the solution of many problems. Fortunately, today mathematical computer programs like *Maple*, *Mathematica* or *Maxima* help us, but sometimes they also capitulate. Later, we come back to these programs.

One very important rule is the integration by parts. This rule follows from the product rule of differentiation

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

Integrating both side gives

$$\begin{aligned} \int \frac{d(uv)}{dx} dx &= \int u(x)dv + \int v(x)du \\ u(x)v(x) &= \int u(x)dv + \int v(x)du \end{aligned}$$

This is most often written in the following way

$$\int u dv = uv - \int v du \tag{7.4}$$

What does it mean? If we can't solve an integral but can transform it into a product (even by multiplication with 1) we can try to integrate the second multiplicand of the product.

We will exemplify this rule with a simple example. We want to integrate $y = \ln(x)$. For this task we define two new functions $u = \ln(x)$ and $v = x$. Therefore $y = u$ and $dv = dx$. We also get $du / dx = 1/x$ or $du = (1/x)dx$. Now we have all to solve the integral.



$$\int \ln(x) dx = \int u dv = uv - \int v du = x \ln(x) - \int x \frac{1}{x} dx = x \ln(x) - x + c \tag{7.5}$$

Don't forget the integration constant. Beside the Mathematica solution of this problem is shown. Note that math programs use natural logs as the default, hence $\text{Log}[x] = \ln(x)$. Additionally, they don't give the integration constant. Don't forget it!

A second simple example. We want to evaluate $y = x e^x$. In this case we have to choose $u = x$. This gives $du = dx$. We assume also $v = e^x$. This results in $dv = e^x dx$. Now we get

$$\int x e^x dx = \int u dv = x e^x - \int e^x dx = x e^x - e^x + c \tag{7.6}$$

Let's turn to another problem. Look at the next Figure (Fig. 7.1). It shows the change in population size of a typical bacterium. The change in size is proportional to the number of bacteria present. If we have 10000 bacteria they can produce, say, 10000 new bacteria, if we have 1000000 they will produce 1000000 new. Therefore $N_{t+1} = rN_t$. The population raises proportional to the present population size. We can model the process by a difference or by a differential equation of the form

$$\frac{\Delta N(t)}{\Delta t} = rN(t)$$

$$\frac{dN(t)}{dt} = rN(t)$$

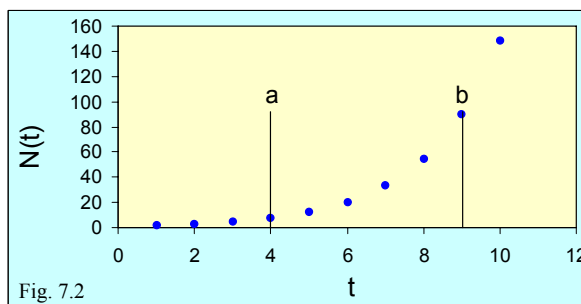
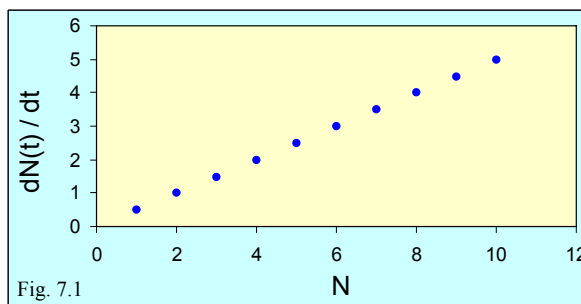
What is the total number of bacteria after time t ? To compute the total number of bacteria N_{total} we first have to solve the so-called **differential equation**

$$\frac{dN(t)}{dt} = rN(t)$$

We multiply both sides with dt , divide through $N(t)$ and integrate

$$\int \frac{1}{N} dN = \int r dt$$

We know already that $d[\ln(x)] / dx = 1 / x$. The derivative of $rt = r$. We get



$$\ln(N) + c_1 = rt + c_2$$

$$N = Ce^{rt}$$

C contains both integrating constants. What is C? If we set t to zero we get $N_0 = C$. That leads to

$$N = N_0 e^{rt}$$

Figure 7.2 shows a plot of N_t against t from $t_0 = a$ (our starting point) to time b. Now we ask about the total number of bacteria that appeared during the time interval from a to b. At time t_0 the number of bacteria was $N_0 = f(t_0) = c$, after some time interval Δt this number is obviously

$$N_1 = f(t_0) + f(t_1) = f(t_0) + f(t_0 + \Delta t)$$

After two time intervals the number will be

$$N_2 = f(t_0) + f(t_1) + f(t_2) = f(t_0) + f(t_0 + \Delta t) + f(t_0 + 2\Delta t)$$

If we divide the whole time interval from a to b into n time intervals Δt we see that the total number of bacteria will be

$$N_{total} = \sum_{i=1}^n f(i)\Delta t$$

This is shown in the next Figure 7.3. We compute the sum of all small rectangles $f(i)\Delta t$. Now we consider very small time differences Δt and use $\lim_{\Delta t \rightarrow 0}$. The Fig. 7.3 shows graphically that the area under the function is the sum of all products $f(t)\Delta t$.

At this stage we remember our indefinite integral. It was

$$F(t) = \int f(t)dt$$

In other words

$$\frac{dF}{dt} = \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} = \frac{d \int f(t)dt}{dt} = f(t) \tag{7.7}$$

or

$$F(t + \Delta t) - F(t) \approx f(t)\Delta t \tag{7.8}$$

We introduce this into our above equation and get

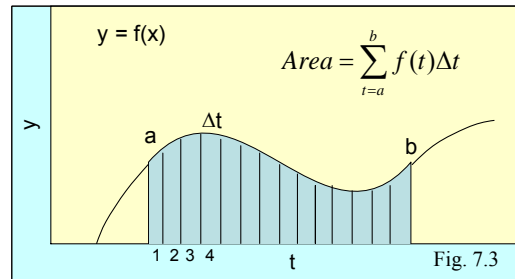
$$N_{total} = \sum_{i=1}^{n-1} [F(i + \Delta i) - F(i)] = \sum_{i=1}^{n-1} F(i + \Delta i) - \sum_{i=1}^{n-1} F(i)$$

Now look again at the Figure 7.3. We denote the time intervals with 1, 2, 3, 4, ... and evaluate the last equation

$$N_{total} = F(2) - F(1) + F(3) - F(2) + F(4) - F(3) \dots + F(n) - F(n-1) = F(n) - F(1)$$

We are able to reduce the whole sum to only two summands, the first and the last.

Now we can rewrite our result and reach in



$$N_{total} = F(b) - F(a) = \int_a^b f(t) dt - \int_a^a f(t) dt \tag{7.9}$$

This is a remarkable result. It tells us that the area under the curve between $x = a$ and $x = b$ is the difference of the integral $F(b) - F(a)$. This is often written in the form

$$\int_a^b f(t) dt = F(t) \Big|_a^b = F(b) - F(a) \tag{7.10}$$

and called the **definite integral** from a to b . The integral sign is in effect a summation sign over very small values dx . In fact, historically it stems from an elongated S for denoting a sum. N_{total} is of course equivalent to the area under the curve from a to b . We can write

$$Area = \lim_{\Delta t \rightarrow 0} \sum_{t=a}^b f(t) \Delta t = \int_a^b f(t) dt \tag{7.11}$$

Now we can solve our initial answer. The number of bacteria is the integral of the accumulation function in the boundaries from a to b . Therefore

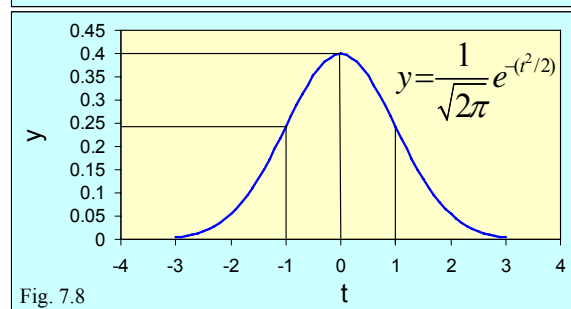
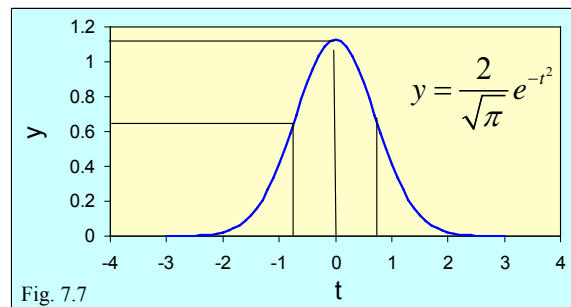
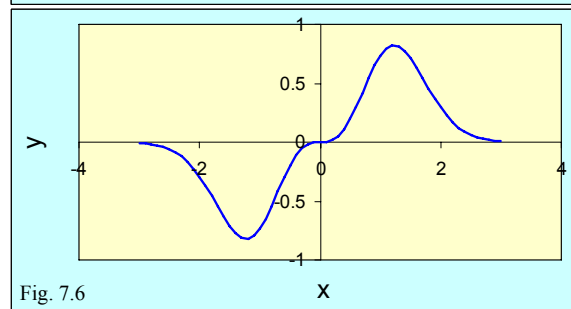
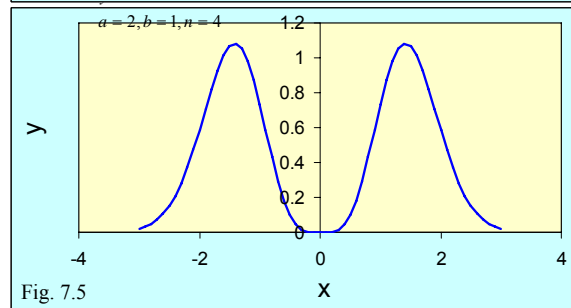
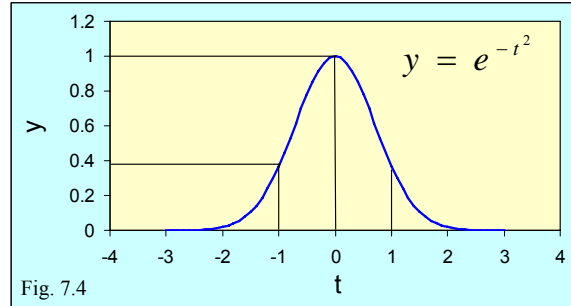
$$N_{total} = \int_a^b N_0 e^{rt} dt = \frac{N_0}{r} e^{rt} \Big|_a^b = \frac{N_0}{r} (e^{rb} - e^{ra}) \tag{7.12}$$

This is the total number of bacteria produced between time a and time b .

Integrating functions is often a difficult task. However, most commonly used integrals are tabulated and today very powerful computer programs like *Mathematica*, *Maple* or *Maxima* are available that solve integrals for us. But this makes things not necessarily easier. We have to interpret the results these programs give us. Very often, complicated integrals do not have exact solutions and the programs give approximations. One very often met approximation is the so-called **error function**. The error function frequently occurs when we try to integrate functions of the following general type (Fig. 7.4)

$$y = f(x) e^{-g(x)^2}$$

Let's consider a simple example. The following function has to be evaluated



$$y = \int ax^n e^{-bx^2} dx$$

Beside (Figs. 7.5 and 7.6) are shown two plots with different values of a, b, and n. *Mathematica* gives us the following output.

$$\frac{a\sqrt{\pi}x^n \operatorname{Erf}[\sqrt{b}] \operatorname{Log}(e)}{2\sqrt{b} \operatorname{Log}(e)}$$

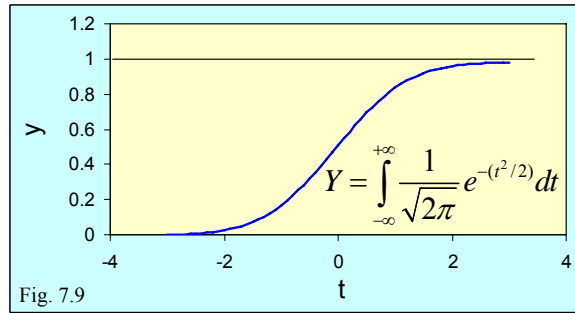


Fig. 7.9

What does this mean? First of all, we can simplify the result by using natural logarithms. $\operatorname{Log}(e)$ becomes 1. Next we detect a general solution and the error function. Separating them gives

$$F(x) = \frac{a\sqrt{\pi}x^n}{2\sqrt{b}} \operatorname{Erf}[\sqrt{b}]$$

What if $\operatorname{Erf}(b^{1/2})$?

The error function is defined for all x by

$$\operatorname{Erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \tag{7.13}$$

It is the solution of the so-called **Gaussian integral** with which we later deal in detail. This integral has the general form (Fig. 7.7)

$$y = \frac{2}{\sqrt{\pi}} e^{-t^2} \tag{7.14}$$

and plays a very important role in statistics. It is often given in a slightly different form (Fig. 7.8)

$$y = \frac{1}{\sqrt{2\pi}} e^{-(t^2/2)} \tag{7.15}$$

It has to be solved numerically or again by an approximation. Figures 7.4, 7.7, and 7.8 show the Gaussian function in three different ways. This function has some important properties. The maximum of the function has the value 1 (or $1 / (2\pi)^{0.5}$ or $2 / \pi^{0.5}$ dependent of the form you choose) at point $t = 0$.

Two other important points are the **points of inflection**. There, the slope of the function has its maximum or minimum. These points are at $t = \pm 1$ (or $t = \pm\sqrt{(1/2)}$). How to compute them? To compute the points of inflection we need the second derivative. This has to be zero. The second derivative of the Gaussian function is

$$y'' = \left(\frac{2}{\sqrt{\pi}} e^{-t^2} \right)'' = \left(\frac{-4t}{\sqrt{\pi}} e^{-t^2} \right)' = \frac{-4}{\sqrt{\pi}} e^{-t^2} + \frac{8t^2}{\sqrt{\pi}} e^{-t^2} = 4e^{-t^2} \left(\frac{t^2 - 1}{\sqrt{\pi}} \right)$$

This function is zero at $t = -1$ and $t = 1$ as shown in the Figures 7.4 and 7.7 .

However, we are more interested in the integral of this function because this is the error function we needed. The next Figure 7.9 shows this function in its most often used form. It shows that the value of Y goes asymptotically to 1. In other words the whole area under the Gauss curve is exactly 1.

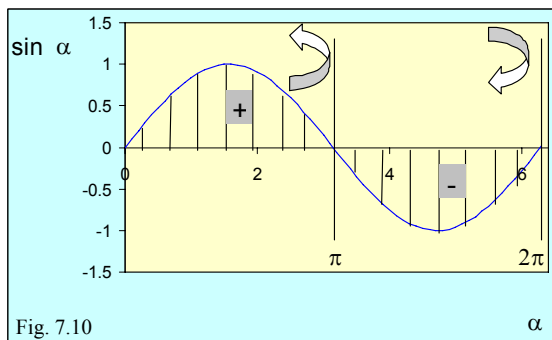
$$Y = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-(t^2/2)} dt = 2 \lim_{x \rightarrow \infty} \int_0^x \frac{1}{\sqrt{2\pi}} e^{-(t^2/2)} dt = 1 \tag{7.16}$$

Additionally holds

$$Y = \int_0^{\infty} e^{-t^2} dt = \frac{\sqrt{\pi}}{2} \tag{7.17}$$

A next example. Acoustics relies on harmonic oscillations. These can be described by trigonometric functions. We want to compute the integral of $y = \sin(x)$ from 0 to 2π . This is the area under the curve as shown beside (Fig. 7.10). *Mathematica* gives us a simple answer.

$$A = \int_0^{2\pi} \sin(x) dx = 0$$



What does this mean? Obviously there is an area under the curve. But one time this area is above the x-axis and one time below. Because the function is symmetrically, the sum of both parts is zero. This leads us to an important thing. It is important how we reached a certain data point. If we reach it in a anti clockwise manner we get a positive area and this is the so-called **mathematical positive sense**. A clockwise manner leads to a **mathematical negative sense**. Therefore, we have to be carefully when relying on numerical computations when the function considered changes its direction (mathematically has positive and negative parts). To solve our initial question we must double the positive part and solve the equation

$$A = 2 \int_0^{\pi} \sin(x) dx$$

We evaluate

$$\frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{\sin(x + \Delta x) - \sin(x)}{\Delta x}$$

A look in a textbook (or some minutes of calculations) give us the following identity

$$\sin(a) - \sin(b) = 2 \sin\left(\frac{a-b}{2}\right) \cos\left(\frac{a+b}{2}\right) \tag{7.18}$$

We get therefore

$$\frac{dy}{dx} = \lim_{\Delta x \rightarrow 0} \frac{\sin(x + \Delta x) - \sin(x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{2}{\Delta x} \sin\left(\frac{\Delta x}{2}\right) \cos\left(\frac{2x + \Delta x}{2}\right) = \lim_{\Delta x \rightarrow 0} \cos(x) \frac{\sin(\Delta x/2)}{\Delta x/2} = \cos(x)$$

$$\frac{d \sin(x)}{dx} = \cos(x) \tag{7.19}$$

The slope of the function $y = \sin(x)$ is for every point x exactly $y' = \cos(x)$, a surprising result. Of course, the latter result could only be obtained if

$$\lim_{\Delta x \rightarrow 0} \frac{\sin(\Delta x/2)}{\Delta x/2} = 1$$

How to prove this. We use again the rule of l'Hospital and assume simply that $\sin'(x) = \cos(x)$. There-

fore (with $\Delta x/2 = h$)

$$\lim_{h \rightarrow 0} \frac{\sin(h)}{h} = \lim_{h \rightarrow 0} \frac{\sin'(h)}{(h)'} = \lim_{h \rightarrow 0} \frac{\cos(h)}{1} = 1$$

Our result implies of course two further simple relations

$$\begin{aligned} \frac{d \cos(x)}{dx} &= -\sin(x) \\ \frac{d \tan(x)}{dx} &= \frac{1}{\cos^2(x)} \end{aligned} \tag{7.20}$$

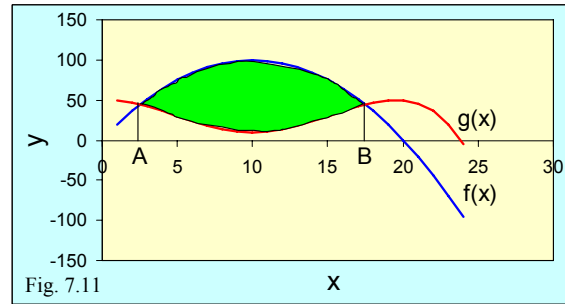


Fig. 7.11

Now, we can solve the initial question about the integral of $\sin(x)$ from 0 to 2π . It is

$$A = 2 \int_0^\pi \sin(x) dx = 2(-\cos(\pi) + \cos(0)) = 4$$

The area under the curve of our trigonometric function is a natural number.

A last example. We are interested in the area beneath the two curves shown in the Fig. 7.11 in green colour. We need the integral of $f(x)$ and $g(x)$ in the boundaries A and B. The area is then the difference between both integrals. Therefore, we have to calculate $F(x)$ and $G(x)$ and take the difference $A = [(F(B) - F(A)] - [(G(B) - G(A)]$. We get

$$A = \int_{x=A}^B f(x) dx - \int_{x=A}^B g(x) dx = \int_{x=A}^B f(x) - g(x) dx$$

This may be written slightly different.

$$A = \int_A^B f(x) - g(x) dx = \int_A^B y \Big|_{g(x)}^{f(x)} dx = \int_A^B \int_{g(x)}^{f(x)} h(x, y) dx dy \tag{7.21}$$

This latter equation is a general form of a so-called **double integral**. Double integrals are used to compute areas bounded by two functions. We see that a simple integral appears as a special case of such an integral were the function $y = g(x) = 0$.

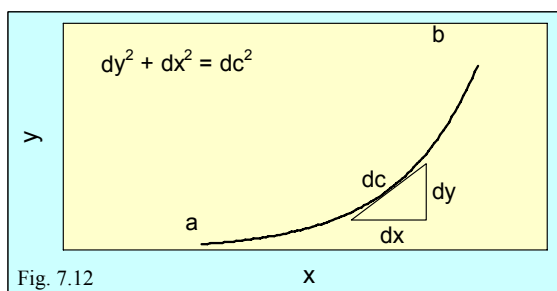
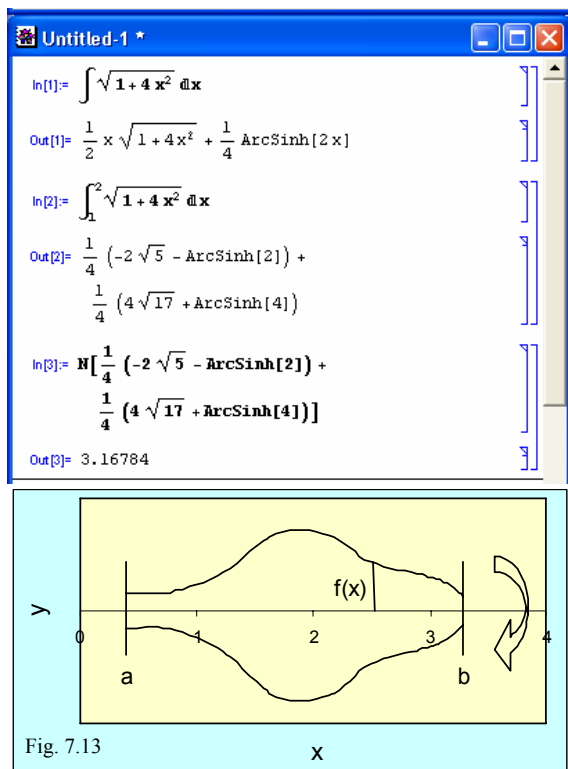


Fig. 7.12

Sometimes it is important to compute the total length of a line. For instance, we follow the troops of an animal and are interested in the total way that animal run or we need to calculate the total length of blood vessels or parts of the xylem system. These are special cases of a general problem to compute the length of a given function $f(x)$. Look at Fig. 7.12. We want to calculate the length of the function

$y=f(x)$ between the points a and b. For that task we apply the law of Pythagoras and assume that the total length is divided into many small parts dc . The total length is then the sum of all dc . Therefore

$$L = \lim_{dc \rightarrow 0} \sum_a^b dc = \int_a^b dc = \int_a^b \sqrt{dy^2 + dx^2} = \int_a^b \sqrt{\frac{dy^2 + dx^2}{dx^2}} dx = \int_a^b \sqrt{1 + f'(x)^2} dx \tag{7.22}$$



An example. What is the length of the function $y = x^2$ from $x = 1$ to $x = 2$? We need

$$L = \int_1^2 \sqrt{1 + 4x^2} dx$$

We apply the math program. The integral is a quite complicated term. Hence we use the numerical solution and get $L = 3.167$.

Another problem concerns the volume of bodies. If these bodies were formed through a rotating process (Fig. 7.13) (they are **rotation bodies**) the computation is easy. Assume the body rotates around the x -axis. Its area at point x would then be equal to $2\pi r^2$, where r is exactly the function value at x , it est $f(x)$. A half rotation gives the full body and we get

$$V = \frac{1}{2} \int_{x_1}^{x_2} 2\pi r f(x) dx = \pi \int_{x_1}^{x_2} f(x) * f(x) dx = \pi \int_{x_1}^{x_2} f(x)^2 dx \quad (7.23)$$

If we rotate around the y -axis we get of course

$$V = \frac{1}{2} \int_{y_1}^{y_2} 2\pi r f(y) dy = \pi \int_{y_1}^{y_2} f(y) * f(y) dy = \pi \int_{y_1}^{y_2} f(y)^2 dy \quad (7.24)$$

Hence we have to take the inverse function $x=f(y)$.

An example. What is the volume of the body generated by the rotation of $y = x^2$ from $x = 1$ to $x = 2$ around the y -axis. We need the inverse function.

$$V = \pi \int_1^4 (\sqrt{y})^2 dy = \frac{\pi}{2} y^2 \Big|_1^4 = 7.5\pi = 23.56$$

At last we are interested in the surface of a rotation body. We need the surface of all small circles with radius r and length ΔL . This surface is $2\pi r \Delta L$. r equals $y = f(x)$ and ΔL is again $\sqrt{(1 + f'(x)^2)}$. Hence

$$M = 2\pi \int_{x_1}^{x_2} y \sqrt{1 + (y')^2} dx \quad (7.25)$$

For a rotation around the y -axis we get

$$M = 2\pi \int_{y_1}^{y_2} x \sqrt{1 + \left(\frac{1}{y'}\right)^2} dy \quad (7.26)$$

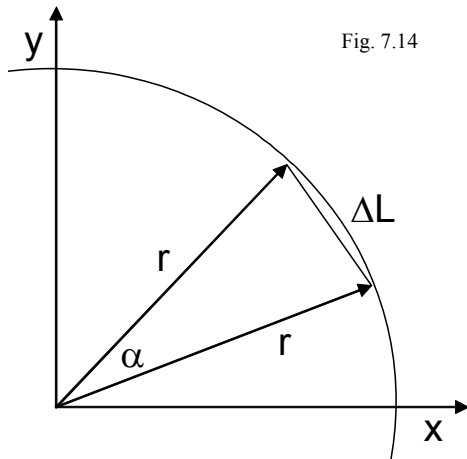


Fig. 7.14

Often it is easier to apply polar coordinates to compute length, areas or volume. To compute the length L of a segment given by the polar function $r = f(\alpha)$ we would have to add up all segments ΔL . If ΔL goes to zero $\Delta L / r$ approaches $\sin(\Delta\alpha)$. Hence (Fig. 7.14)

$$L \approx \sum \Delta L = \sum r \sin(\Delta\alpha)$$

Now we use a second approximation. We look for the limes of $\sin(x) / x$. We use *Mathematica* and compute the lim directly and via a Taylor series expansion of $\sin(x)$. In the case of the Taylor series the terms higher than the first are small in relation to the first and

```

Untitled-1 *
In[3]:= Limit[Sin[x] / x, x -> 0]
Out[3]= 1
In[10]:= Log1 == Series[Sin[x], {x, 0, 5}]
Out[10]= Log1 == x - x^3/6 + x^5/120 + O[x]^6
    
```

we drop them. Both approaches result therefore in

$$\lim_{x \rightarrow 0} \sin(x) = x$$

Now we introduce this approximation into the above equation and get

$$L \approx \sum r \sin(\Delta\alpha) \approx \sum r \Delta\alpha = \int r d\alpha \tag{7.27}$$

For example, what is the circumference of a circle? We consider the range of $\alpha = 0$ to $\pi / 2$. Hence

$$L = 4 \int_0^{\pi/2} r d\alpha = 4r\alpha \Big|_0^{\pi/2} = 2\pi r$$

Another example. What is the length of the logarithmic spiral in Fig. 7.17 from 0 to 2π ? The length is

$$r = c^\alpha \rightarrow L = \int_0^{2\pi} c^\alpha d\alpha = \frac{c^\alpha}{\ln(c)} \Big|_0^{2\pi} = \frac{c^{2\pi} - 1}{\ln(c)}$$

By a similar reasoning we get the area inside the range of $\Delta\alpha$. This area is approximately $r \cdot r \cdot \sin(\alpha) / 2$.

We get

$$A = \frac{1}{2} \sum r^2 \sin(\Delta\alpha) \approx \frac{1}{2} \sum r^2 \Delta\alpha = \frac{1}{2} \int_{\alpha_1}^{\alpha_2} r^2 d\alpha \tag{7.28}$$

Now it is easy to compute the area of a circle. It is

$$A = 4 \frac{1}{2} \int_0^{\pi/2} r^2 d\alpha = 2r^2 \alpha \Big|_0^{\pi/2} = \pi r^2$$

Literature

(Coloured titles are available in the Institute or the library, red titles are of major importance)

Cornish Bowden A. 1999—Basic Mathematics for Biochemists—Oxford Univ. Press 2nd. Ed.

Ennos R. 1999—Statistical and Data Handling Skills in Biology—Longman.

Foster P. C. 1998—Easy Mathematics for Biologists— Taylor and Francis.

Jordan D. W., Smith P. 2002. Mathematical techniques. 3rd. Ed. Oxford Univ. Press

Grossman S., Tuner J. E. 1974—Mathematics for the Biological Sciences—Macmillan.

Martin J. 1972—Podstawy Matematyki i Statystyki—Warszawa.

Portenier C., Gromes W. 2003. Mathematik für Biologen und Humanbiologen. Script. Marburg

Scheiner S. M., Gurevitch J. (eds.) 2001—Design and Analysis of Ecological Experiments— Oxford Univ. Press (2nd. Ed.).

Wilson E. O., Bossert W. H. 1971, A Primer of Population Biology. Sinauer (Stamford).

Murray J. D. 2003. Mathematical Biology. 3rd ed. Parts I and II. Springer New York.

Science Magazine special feature. 2004. Mathematics in biology. Science 303.

Napiórkowski K. 2001. Matematyka. http://info.fuw.edu.pl/~ajduk/FUW/matnkf/matematyka01_nkf.pdf

Online archives and textbooks

Online Mathematical textbooks (A large collection of textbooks) <http://www.math.gatech.edu/~cain/textbooks/onlinebooks.html>

General mathematics (a collection of online lecture scripts and basic text on mathematics) http://www.geocities.com/alex_stef/mylist.html

Mathematics online (a source of educational online texts) <http://www.glencoe.com/sec/math/>

Mathematics Virtual Library (Many links to interesting web pages and programs) <http://www.math.fsu.edu/Science/math.html>

Math on the web (Search engine for all sorts of mathematics) <http://www.ams.org/mathweb/mi-mathinfo07.html>

The Math Archive (Many links to interesting web pages and programs) <http://archives.math.utk.edu/>

Eric Weisstein's Mathematics (a large online mathematics dictionary, with many examples) <http://mathworld.wolfram.com/>

The Internet Mathematics library (a large collections of topics for pupils and students, math-beginners) <http://mathforum.org/library/>

Mathematic resources (a large compilation of math internet pages) <http://www.clifton.k12.nj.us/cliftonhs/chsmedia/chsmath.html>

Kolegium nauczycielski. Materiały z wykładów. (Online scripts on various topics) <http://info.fuw.edu.pl/~ajduk/lect.html>

Johannes Müller. 2003. Mathematical models in biology. Lecture term at TU Munich. http://www-m12.ma.tum.de/lehre/model_2003/skript/skript.pdf

Population growth models (a nice collection of growth models) <http://www.math.duke.edu/education/postcalc/growth/contents.html>.

Population growth models (A collection of growth model an animations) http://members.optusnet.com.au/exponentialist/Growth_Models.htm

Competition models (for persons who are interested In a discussion of the Lotka Volterra models) <http://www.ub.rug.nl/eldoc/dis/fil/r.c.looijen/c11.pdf>

The MacTutor history of mathematics (a very nice page on historical topics) <http://www-history.mcs.st-andrews.ac.uk/>.

Excel Turorials (Many macros) <http://www.herber.de/index.html?http://www.herber.de/forum/archiv/104to108.htm>.

Computational molecular Biology. (a very good side with examples how to use mathematics in molecular biology). <http://www.cs.bc.edu/~clote/ComputationalMolecularBiology/>

Mathematical software

The Windows software collection (public domain and freeware)

<http://archives.math.utk.edu/software/.msdos.directory.html> (contains many very nice programs)

The mathematics virtual library (a collection of software pages) <http://www.math.fsu.edu/Virtual/index.php?f=21>.

Guide to mathematical software (a search engine for math programs) <http://gams.nist.gov/>

Step by step derivatives (a very good program for computing derivatives) <http://www.calc101.com/webMathematica/derivatives.jsp#topdoit>

Derivative calculator (a nice small but quite effective program for computing derivatives) <http://cs.jsu.edu/mcis/faculty/leathrum/Mathlets/derivcalc.html>

JAVA Mathlets for Math Explorations (a nice collection of small math programs for everybody) <http://cs.jsu.edu/mcis/faculty/leathrum/Mathlets/>

The integrator (a small but effective integration program)
<http://www.integrals.com/index.en.cgi>

The MathServ Calculus toolkit (a collection of Math applets for calculus computation)
<http://www.math.vanderbilt.edu/~pscrooke/toolkit.html>

Modelowanie rzeczywistości (a nice Polish page with a program collection and many further links) <http://www.wiw.pl/modelowanie/>

Maple homepage. <http://www.maplesoft.com/>

Mathematica homepage (Wofram research) <http://www.wri.com/>

Mathworks homepage (Matlab) <http://www.mathworks.com/>

Mathtype (Office build in tool for mathematics writing) <http://www.mathtype.com/en/products/mathtype/>