Mathematics with DERIVE as Didactical Tool

Wolfram Koepf
University Kassel
`koepf@mathematik.uni-kassel.de`

1   Introduction

In this talk, I would like to show you how DERIVE can be used as a didactical tool in secondary high school education. For my demonstration I use the newest release DERIVE for Windows 5 which supports the use of worksheets. I will concentrate on some examples describing general topics that I find most appropriate to enhance our educational opportunities. In this introduction, I would like to embed the sooner or later expected replacement of numerical calculators in mathematics education by symbolic ones into a historical context.

In the history of mathematics from time to time new calculation tools were invented replacing old ones. Obviously the known mathematical calculation techniques were quite different at the time before logarithm tables made the computation of products and powers easier, and the traditional techniques changed when these methods were generally accessible. Similarly, when the slide-rule came up, the used—and taught—calculation techniques changed again.

More dramatically, when numerical calculators came into the scene, they made both the logarithm tables and the slide-rule dispensable, and, indeed in our days students do no longer have the knowledge about these former mighty tools,[1] let alone their use.

As an example, I have experienced that since the numerical calculator was introduced into the classroom, students tend to avoid working with rational arithmetic. They prefer to use decimal numbers instead. Even if these computations might give inaccurate results in certain situations (e.g. 0.999999999 rather than 1), this is one of the expected results of the use of the new tool, whether we like it or not. The students don't find this an essential disadvantage. Whenever mathematical tools are replaced by new ones, teaching and exercising changes. Obviously we should be aware of this fact, and we should try to make sure that important knowledge and techniques do not get lost.

In my opinion, computer algebra tools are more adequate for mathematics education than numerical calculators are. Whereas a numerical calculator gives some new possibilities like

---

[1]Ask your students whether they can tell you what a slide-rule is!

the fast computation of trigonometric functions etc., it cannot really replace the former computations as the simple example of rational arithmetic shows. Hence, one can expect positive side effects when computer algebra comes into the game.

The examples of this article are mainly taken from my book [3] which offers material for the use of DERIVE in the secondary education at German high schools. I hope I can convince you that save the large personal involvement of each teacher, there are great chances with the use of computer algebra in the classroom.

2   Probability

Computer algebra systems should be used to work with large integers. As already mentioned, since the numerical calculator was introduced into the classroom, students tend to avoid working with rational arithmetic. Rational arithmetic comes back into the picture when using computer algebra.

Here I give an example from probability. The point is that such computations are out of reach for hand computations although they form interesting and by no means pathological questions. It is a typical situation that realistic "real world problems" are not generally accessible for hand computation.

One can use numeric computation for such problems, however the student has no chance to do the necessary error analysis. With some calculators, the result may be even wrong since large factorials are involved.[2] Using exact rational arithmetic, errors do not occur at all. Nevertheless all results presented in this lecture can be computed in a second on today's PC's.

Problem 1:
What is the probability P_1 to throw exactly 50 times Head when throwing a fair coin 100 times?

Clearly, the teacher should develop the theory behind such questions, and the binomial distribution will show up. Hence the result is given by the formula

$$P_1 = \binom{100}{50} \cdot \left(\frac{1}{2}\right)^{100} ,$$

$\binom{n}{k}$ denoting the binomial coefficient. In DERIVE, this can be entered as

```
COMB(100, 50) * (1/2)^100
```

---

[2]Indeed, large factorials are only involved if a numerical calculator does not support products. In a computer algebra system binomial coefficients are computed differently and more efficiently.

and simplification yields

$$P_1 = \frac{12611418068195524166851562157}{158456325028528675187087900672} \, ,$$

a rational number with huge numerator and denominator. Such a result is not accessible with a numerical calculator. On the other hand, in DERIVE we can now approximate the previous exact result and get

$$P_1 = 0.07958923738 \, ,$$

clearly accurate in all its decimal places, and no error analysis is necessary. To interpret the result, it might be unexpected that in almost 8 % of all repetitions of the experiment to throw a coin 100 times *exactly* 50 times Head occurs. Clearly this provokes an even more complicated question like

Problem 2:
What is the probability P_2 to throw between 45 and 55 times Head when throwing a fair coin 100 times?

The binomial distribution gives the sum

$$P_2 = \sum_{k=45}^{55} \binom{100}{k} \cdot \left(\frac{1}{2}\right)^{100} \, ,$$

which is in DERIVE

  SUM(COMB(100, k) * (1/2)^100, k, 45, 55) .

DERIVE computes very fast

$$P_2 = \frac{2886864192022845142126938 9993}{3961408125713216879677197 5168} \, .$$

Clearly, such a computation is absolutely intractable by hand computation! The decimal analogue of the result is

$$P_2 = 0.7287469759 \, .$$

Hence in more than 2/3 of all repetitions of the experiment to throw a coin 100 times we get between 45 and 55 times Head.

This gives a hint how "wide" the underlying binomial distribution is. To measure its width, we compute the standard deviation $\sigma = \sqrt{n\,p\,(1-p)}$,

  $\sigma := \texttt{SQRT}(100 * (1/2) * (1/2))$,

to obtain $\sigma = 5$. This corresponds to the well-known fact that in about 2/3 of all cases the result of a normally-distributed random variable lies in the 1-σ neighborhood $[\mu - \sigma, \mu + \sigma]$ of the expected value μ, which in our case is $\mu = 50$.

3

One could also easily check the 2-σ and 3-σ neighborhoods of the expected value which should contain about 95 % and 99 % of the results, respectively, therefore proving that the binomial distribution is already "almost normal" if $n = 100$.
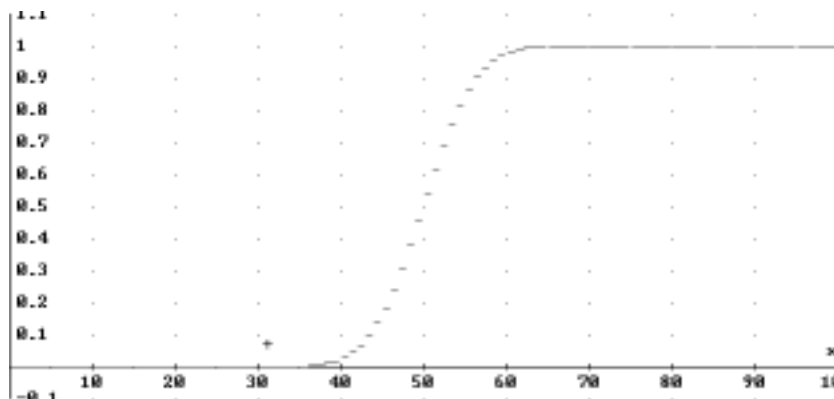
Another hint for the width of the binomial distribution as well as for its similarity to the normal distribution is given by its graph
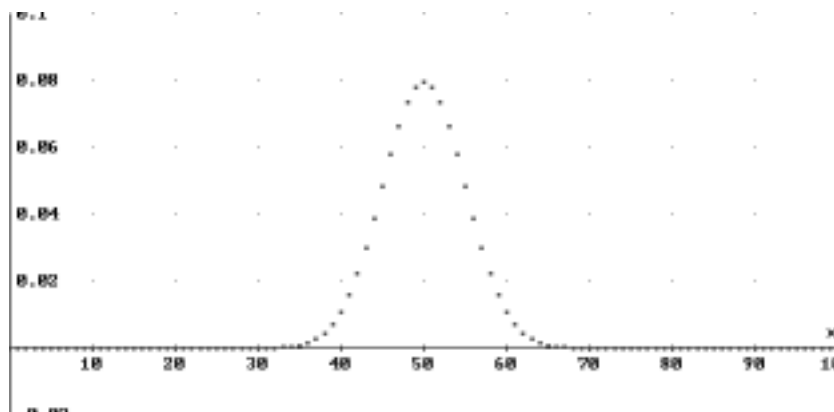
which we receive after simplifying

$$\texttt{VECTOR}([k, \texttt{COMB}(100, k) * (1/2)\texttt{\^{}}100], k, 0, 100)$$

Next, we generate a picture of the distribution function of the binomial distribution

by graphing

$$\texttt{SUM}(\texttt{COMB}(100, k) * (1/2)\texttt{\^{}}100 * \texttt{CHI}(k, x, \infty), k, 0, 100)$$

to be compared with the distribution function of the corresponding Gaussian normal distribution

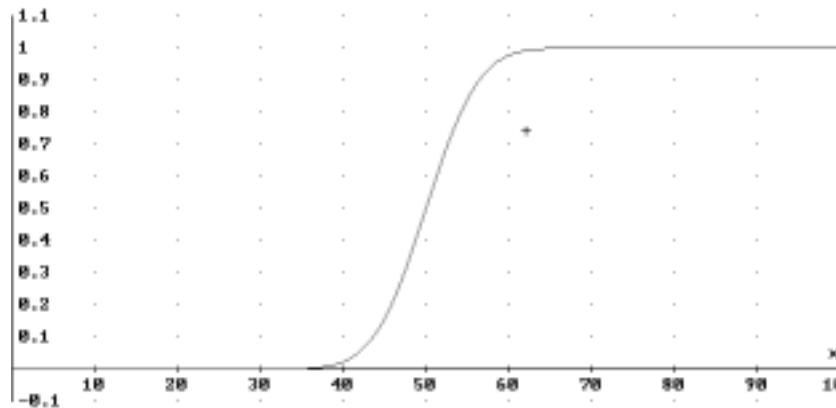$$N(\mu, \sigma, x) := \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \ .$$

A plot of the simplified expression

4

$$\texttt{INT}(1/(\sigma * \texttt{SQRT}(2 * \pi)) * \texttt{EXP}(-(x - 50)\texttt{\^{}}2/(2 * \sigma\texttt{\^{}}2)), -\infty, x)$$
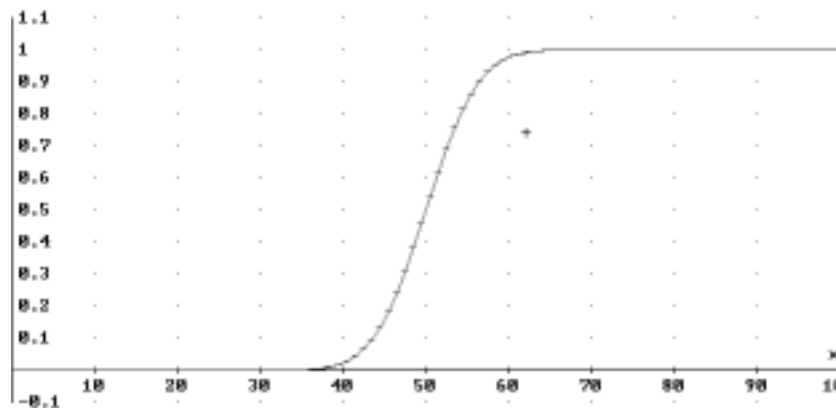
involving the error function

$$\frac{\operatorname{erf}\left(\frac{\sqrt{2}x}{10} - 5\sqrt{2}\right) + 1}{2}$$

yields

The superposition of both functions is most impressive:

Obviously we are free—and DERIVE offers the capability—to study, e.g., also the situation for $n = 1000$.

3   Graphics

We already saw in the above section, that graphs can give very interesting insights into mathematical situations. On the other hand, graphing is not that easy in all situations as the following example shows [4]:
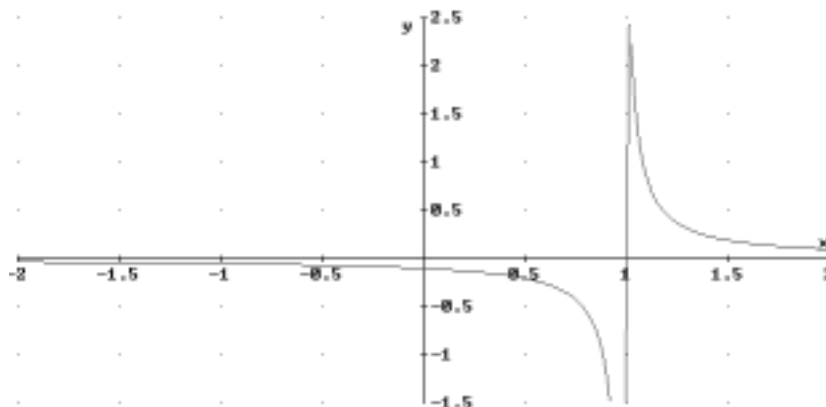
Problem 3:
Graph the rational function $f(x) = \frac{1000\,(x-1)}{(101x-100)\,(100x-99)}$. Where is the second pole?

We enter

$$f := 1000 * (x - 1)/((101 * x - 100) * (100 * x - 99)) \,.$$

DERIVE immediately graphs f as

It looks as if there is a maximum just right of 1, a zero at 1, and a pole just left of 1. Actually, this picture is not bad compared to the results of other computer algebra software. But nevertheless: A glimpse at the formula shows that there should be two poles. Where is the second pole? Even zooming out does not give any hint, so what is going on?

Clearly, directly from the formula we see that the two poles are at $x = \frac{100}{101}$ and $x = \frac{99}{100}$, both slightly smaller than 1. To find out how the graph between the two poles looks, we search for the critical points by

`SOLVE(DIF(f, x), x)`

and get the result

$$x = \pm\infty \quad \text{or} \quad x = 1 - \frac{\sqrt{101}}{1010} \quad \text{or} \quad x = \frac{\sqrt{101}}{1010} + 1 \,.$$

The last solution is obviously larger than 1, hence yields the maximum, so the middle solution is the one we are interested in. Substituting this number into f

`SUBST(f, x, 1 − SQRT(101)/1010)`

yields the value $20000\sqrt{101} + 201000 \approx 4.019975124 \cdot 10^5$. Now we see what's happening: The function has a local minimum at $x = 1 - \frac{\sqrt{101}}{1010}$ and all its values are larger than $4 \cdot 10^5$ between the two poles. That's why we didn't see anything! Using a reasonable plot range[3] gives the following graph:

---

[3]using $\boxed{\texttt{Set: Plot Range}}$ in the 2D-Plot window
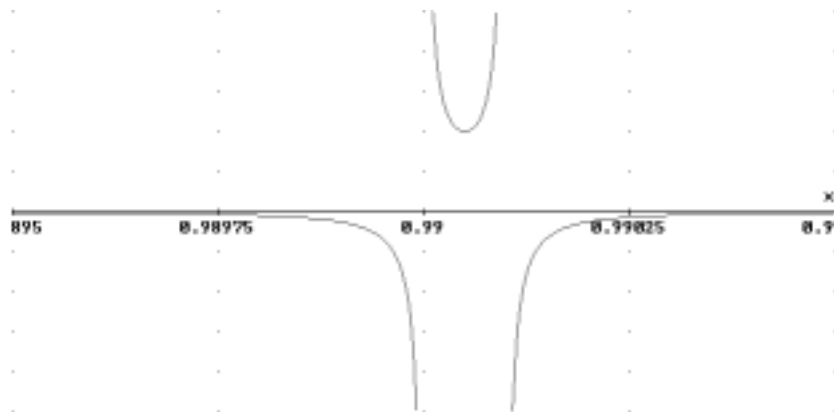
This example shows that the use of computer algebra enhanced graphing is much superior to the use of a graphics calculator with which such information is not available.

4   Factorization

In this section we deal with a question which is out of reach for hand computation as well as for graphics calculators. Here, symbolic computation is the essential ingredient.

Leibniz, one of the developers of the differential and integral calculus, still doubted the existence of an elementary antiderivative of the rational function

$$\frac{1}{1+x^4} \ .$$

The reason was that he was not able to find a proper *real factorization* of the denominator polynomial

$$1 + x^4$$

in terms of quadratics. Let's check this with DERIVE. We enter

$g := 1/(1 + x^4)$

and use DERIVE to factorize g by

$\texttt{FACTOR}(g)$

to get

$$1 + x^4$$

again. This seems to approve Leibniz's opinion. However, we get more detailed information when using DERIVE's $\boxed{\texttt{Simplify}}$ menu. The $\boxed{\texttt{Simplify: Factor}}$ submenu asks the user for the degree of factorization: Trivial, Rational, Radical or Complex, Rational being

7

the default choice. This gives us a clue why a proper factorization was not found in the first step: A factorization of g over the rationals $\mathbb{Q}$ does not exist![4] However, since we search for a factorization over $\mathbb{R}$, we choose the Radical option to get

$$1 + x^4 = (x^2 + \sqrt{2}\,x + 1)\,(x^2 - \sqrt{2}\,x + 1)\,.$$

With this factorization in mind, it is easy to find the partial fraction decomposition

$$\frac{1}{1+x^4} = -\frac{\sqrt{2}\,x}{4\,(x^2 - \sqrt{2}\,x + 1)} + \frac{1}{2\,(x^2 - \sqrt{2}\,x + 1)} + \frac{\sqrt{2}\,x}{4\,(x^2 + \sqrt{2}\,x + 1)} + \frac{1}{2\,(x^2 + \sqrt{2}\,x + 1)}$$

with $\boxed{\texttt{Simplify:\ \ Expand}}$, and to understand the antiderivative

$$\int \frac{1}{1+x^4}\,dx = \frac{\sqrt{2}\,\arctan\left(\sqrt{2}\,x - 1\right)}{4} + \frac{\sqrt{2}\,\arctan\left(\sqrt{2}\,x + 1\right)}{4} - \frac{\sqrt{2}\,\ln\left(\frac{x^2 - \sqrt{2}\,x + 1}{x^2 + \sqrt{2}\,x + 1}\right)}{8}$$

which is returned by DERIVE's integrator.

5   Parametric Representation of Conic Sections

Students of Physics need the parametric representation of conic sections when planetary motion is studied. University teachers in Germany observe and criticize that a typical student does not have this knowledge. The reason is simple: With hand computations the parametric representation of conic sections is rather difficult to derive. Here, we show how a rational factorization gives this result very easily. Note that although rational factorizations are difficult to find, they are very easy to check.

We start with the rectangular equation of an ellipse whose left focus lies in the origin:

$$(x - e)^2/a^2 + y^2/b^2 - 1 = 0\,.$$

We replace b^2 by $a^2 - e^2$ using the substitution

$$(x - e)^2/a^2 + y^2/\texttt{SQRT}(a^2 - e^2)^2 - 1 = 0\,.$$

Simplification yields

$$\frac{x^2\,(a^2 - e^2) + 2\,e\,x\,(e^2 - a^2) + a^2\,y^2 - (a^2 - e^2)^2}{a^2\,(a^2 - e^2)} = 0\,.$$

---

[4]Actually this was *proved* by DERIVE in the first calculation!

Next, we introduce polar coordinates $x = r \cos \phi$ and $y = r \sin \phi$ by another substitution. Simplification gives

$$\texttt{equation} := \frac{e^2 \, r^2 \, \texttt{COS}(\phi)^2 + 2 \, e \, r \, (a^2 - e^2) \, \texttt{COS}(\phi) + a^4 - a^2 \, (2 \, e^2 + r^2) + e^4}{a^2 \, (e^2 - a^2)} = 0 \; .$$

You see that the simplification procedure has removed the sine expression using the trigonometric identity $\sin^2 \phi = 1 - \cos^2 \phi$. But what's next?

What we want is a representation of r in terms of ϕ. It turns out that students who don't have any idea how to proceed, but are accustomed with DERIVE, nevertheless have the idea to let DERIVE solve the equation for r, and this is a successful approach! On the other hand the main reason for this success is hidden. Let's rather try to factorize the given expression by

```
FACTOR(equation)
```

with the result

$$\frac{(e \, r \, \texttt{COS}(\phi) + a^2 + a \, r - e^2) \, (e \, r \, \texttt{COS}(\phi) + a^2 - a \, r - e^2)}{a^2 \, (a + e) \, (e - a)} = 0 \; .$$

It is now obvious why the equation easily can be solved for r: By the factorization, either of the two numerator factors must be zero. Using

```
SOLVE(equation, r)
```

returns these two solutions

$$r = \frac{e^2 - a^2}{e \, \texttt{COS}(\phi) - a} \quad \text{or} \quad r = \frac{e^2 - a^2}{e \, \texttt{COS}(\phi) + a} \; .$$

It turns out that only one of these is positive, hence we get

$$r = \frac{e^2 - a^2}{e \, \texttt{COS}(\phi) - a} \; .$$

Finally, we introduce the eccentricity $\varepsilon = e/a$ which is smaller than 1 for an ellipse, larger than 1 for a hyperbola and equal to 1 for a parabola, and obtain

$$r = \frac{e \, (\varepsilon^2 - 1)}{\varepsilon \, (\varepsilon \, \texttt{COS}(\phi) - 1)} \; .$$

It turns out ([3], Chapter 2) that a similar computation yields the same result for a hyperbola, for which, however, $\varepsilon > 1$.

9

To get a picture, we produce a graph for $e = 1$, $\varepsilon = 0.9$ and $\varepsilon = 1.1$, respectively.[5] This yields
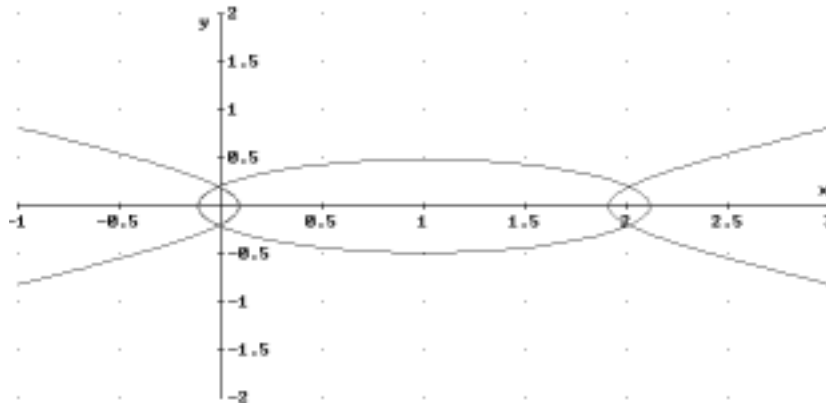
6   Cubic Newton Iteration

Let's start to derive the usual Newton iteration by linearization. Declare f by

$$f(x) :=$$

to be an arbitrary function of a single variable. Linearizing f at the point x_n gives the equation

$$y = f(\texttt{xn}) + (x - \texttt{xn}) * f'(\texttt{xn}) .$$

If x_n is an approximation of a zero, we therefore can set $y = 0$, and solve for x to get the next iteration value:

$$\texttt{SOLVE}(0 = f(\texttt{xn}) + (x - \texttt{xn}) * f'(\texttt{xn}), x)$$

with the solution

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

This gives the well-known Newton iteration formula for x_{n+1} which usually converges quadratic to a simple zero of f (see e.g. [2] for details).

Now we are ready to develop a cubic interpolation scheme (see e.g. [3], Chapter 3) by starting with a second order parabolic approximation of f:[6]

---

[5]To generate a polar plot, one uses $\boxed{\texttt{Set: Coordinate System: Polar}}$ in the 2D-Plot window.
[6]Instead of using the TAYLOR function, one could also enter the expression by hand.

$$y = \texttt{TAYLOR}(f(x), x, \texttt{xn}, 2)$$

which simplifies to

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

Since $x - x_n \approx -\frac{f(x_n)}{f'(x_n)}$ by the Newton iteration, one can linearize the quadratic factor. Setting $y = 0$ and solving for x, again, we get

$$\texttt{equation} := \texttt{SOLVE}(0 = (x - \texttt{xn}) * (-f'(\texttt{xn})/f(\texttt{xn}))/2 + (x - \texttt{xn}) * f'(\texttt{xn}) + f(\texttt{xn}), x)$$

with the solution
$$x = \frac{2 \, f(x_n) \, f'(x_n)}{f(x_n) \, f''(x_n) - 2 \, f'(x_n)^2} + x_n \ .$$

This is a cubic iteration scheme as we shall see. Let's test it to calculate square roots. To compute $\sqrt{a}$, we use

$$f(x) := x^2 - a \ .$$

With this definition, one can simplify $\texttt{equation}$ again to obtain

$$x = 3 \, x_n - \frac{8 \, x_n^3}{a + 3 \, x_n^2} \ .$$

Let's check that this gives a cubic iteration scheme: Simplifying $x_{n+1} - a$ by

$$\texttt{FACTOR}((3 * \texttt{xn} - 8 * \texttt{xn}^3/(a + 3 * \texttt{xn}^2))^2 - a)$$

gives
$$\frac{(x_n^2 - a)^3}{(a + 3 \, x_n^2)^2}$$

proving the cubic convergence. The same method shows that Newton's iteration is quadratic convergent.

Now, we test the iteration for

$$a := 10 \ .$$

The iteration can be executed by the command

`ITERATES`$(3 * \mathtt{xn} - 8 * \mathtt{xn}^3/(a + 3 * \mathtt{xn}^2), \mathtt{xn}, 1, 5)$

which we approximate to get

$$[1, 2.384615384, 3.144897306, 3.162277527, 3.16227766, 3.16227766].$$

Hence $\sqrt{10}$ is computed from the (bad) initial guess 1 in 4 steps, whereas Newton's iteration

`ITERATES`$(\mathtt{xn} - f(\mathtt{xn})/f'(\mathtt{xn}), \mathtt{xn}, 1, 5)$

gives
$$[1, 5.5, 3.659090909, 3.196005081, 3.162455622, 3.162277665]$$
convergence in five steps. From the practical point of view this might be not very essential, but from a theoretical point of view the higher order Newton method is rather interesting.

Note that by a similar computation iteration schemes with convergence of arbitrary order can be obtained which however—as a negative computational repercussion–get more and more complicated. Note further that the cubic iteration scheme for the square root discussed in this section was discovered by Dedekind (see e.g. [5], pp. 99).

7   Final Remarks

I hope that I could convince you that the smart use of computer algebra in the classroom can be a very convenient didactical tool to teach and develop mathematical insights. On the other hand, one should not disregard the fact that nevertheless the students should obtain advanced mathematical skills before using computer algebra since otherwise they might neither be able to handle such a program nor to understand its results. A similar remark applies to hand-held computer algebra tools like the TI-89. Therefore, it is very important that the essential mathematical skills are still practiced.

Especially if you have weak students, it may happen that this scheme cannot be put into practice. One essential problem are time constraints. Furthermore, my experience is that weaker students tend to be overwhelmed by the handling of a computer algebra system[7] since they need to acquire even more skills to use such a program which they try to avoid.

Note that in the meantime many different teachers all over the world use computer algebra in Math education and curriculum discussions are on their way. A collection of interesting projects on this topic can be found in Section 3.8 on "Computer Algebra in Education" which I compiled for [1].

---

[7]even though the students might enjoy the calculation assistance

References

[1] Grabmeier, J., Kaltofen, E. and Weispfenning, V. (Eds): *New Reference Book on Computer Algebra*, Springer, Berlin, 2000.

[2] Koepf, W., Ben-Israel, A. and Gilbert, R. P.: *Mathematik mit Derive*, Vieweg, Braunschweig/Wiesbaden, 1993.

[3] Koepf, W.: *Derive für den Mathematikunterricht*, Vieweg, Braunschweig/Wiesbaden, 1996.

[4] Koepf, W.: Numeric versus symbolic computation. International Journal of Computer Algebra in Mathematics Education **5** (1), 1998, 29–54.

[5] Schmersau, D. and Koepf, W.: *Die reellen Zahlen als Fundament und Baustein der Analysis.* Oldenbourg, Munich, 2000.