Summary 1 – Part 3

Symbolic and numerical solution of equations and systems of equations, comments and additions to differential equations

Michel Beaudin, ETSMTL, Montréal, Kanada

2 Fixed Point Method and Newton's Method

Let's start with a scalar equation of the type $g(x) = x$ to introduce the Banach contraction principle.

2.1 Theorem Let g a continuous function on the closed interval I, such that $g(x) \in I \ \forall x \in I$.

a) Then g has (at least) one fixed point r, i.e. $\exists\ r \in I$ such that $g(r) = r$.

b) Suppose further that g is derivable on the interior of I and that there exists a constant K such that $|g'(x)| \leq K < 1 \ \forall x \in I$. Then the equation $g(x) = x$ has a unique solution in I. Moreover, the sequence of successive approximations defined by $x_{n+1} = g(x_n),\ n = 0, 1, 2, \ldots$ converges to this solution, regardless of the initial point $x_0 \in I$.

(The reason we talk about "contraction" is that, by the mean value theorem, we have, if x and $y \in I$, that there is a number z such that $g(x) - g(y) = g'(z)(x - y)$, hence $g(x) - g(y)| \leq K|x - y|$. Since K is less than 1 this means that g "brings" the points together, so g is a "contraction").

Proof:

a) by sketching a graph, we "prove" part a) since if since if $I = [a, b]$ and if $g(a) = a$ and/or if $g(b) = b$, we already have (at least) one fixed point. By assumption the image of I is contained in I and if $g(a) > a$ and $g(b) < b$ then the continuity of g (or the intermediate value theorem applied to the function $h(x) = x - g(x)$ give the result.

b) For uniqueness, note that if we have two fixed points, say x and y, then the mean value theorem implies the existence of a z between x and y such that $g(x) - g(y) = g'(z)(x - y)$. Thus

$$|x-y| = |g(x)-g(y)| = |g'(z)(x-y)| = |g'(z)||x-y| \leq K|x-y|$$

hence $|x-y|(1-K) \leq 0$. With $1 - K > 0$ follows that $x = y$. Let's move on to existence. Note that if we define a sequence by $x_{n+1} = g(x_n),\ n = 0, 1, 2, \ldots$ where $x_0 \in I$ arbitrary, then this sequence is well defined since the image of g remains in I. Moreover, if this sequence converges to a number say r, then this number is necessarily a fixed point of g. Indeed, by two theorems of analysis, the limit of a convergent sequence of elements of a closed interval belongs to this interval and a continuous function "carries" convergent sequences, so

$$r = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} g(x_n) = g\left(\lim_{n\to\infty} x_n\right) = g(r).$$

It thus remains to show that the sequence, defined by $x_{n+1} = g(x_n),\ n = 0, 1, 2, \ldots$, is convergent. We note that $x_n = x_0 + (x_1 - x_0) + (x_2 - x_1) + \cdots \quad - x_{n-1})$. Therefore, the sequence will $\{x_n\}_{n=0}^{\infty}$ converge if and only if the series $\sum_{j=0}^{\infty}(x_{j+1} - x_j)$ converges. To demonstrate this, the contraction hypothesis is used and the "good old" geometric series as well. Indeed:

$$|x_2 - x_1| = |g(x_1) - g(x_0)| \le K|x_1 - x_0|$$

$$|x_3 - x_2| = |g(x_2) - g(x_1)| \le K|x_2 - x_1| \le K^2|x_1 - x_0|$$

$$\vdots$$

$$|x_{j+1} - x_j| \le K^j|x_1 - x_0|$$

Since $K < 1$, the geometric series $|x_1 - x_0|\sum_{j=0}^{\infty} K^j$ is convergent and the comparison criterion implies the convergence of the series $\sum_{j=0}^{\infty}|x_{j+1} - x_j|$. And because absolute convergence implies convergence, the proof is complete. ♦

2.2 Examples: Several details will be given in class for the following examples.

2.2.1 The equation $x = \frac{x}{2} + \frac{1}{x}$ can easily be solved in its equivalent form $x^2 = 2$. With $g(x) = \frac{x}{2} + \frac{1}{x}$ and $x_0 = 1$, the solution converges towards $1.4142136 \approx \sqrt{2}$.

2.2.2 The equation $x^3 - 3x + 1 = 0$ has three real and distinct roots (we have shown how to find them in exact mode in 1.6), one of which is between 1 and 2. We will see that $g(x) = \frac{x^3+1}{3}$ or $g(x) = \frac{1}{3 - x^2}$ do not work but that the procedure works applied to $g(x) = \frac{3}{x} - \frac{1}{x^2}$ (in other words, this function satisfies the assumptions of Theorem 2.1b) and we find 1.53209).

2.2.3 The equation $x = 2\cos x$ has one solution between 0 and $\pi/2$. While $g(x) = 2\cos x$ doesn't work, $g(x) = \frac{x + 2\cos x}{2}$ is working pretty well (one finds 1.02987). We will show here how *Derive* and its function FIXED_POINT(f, x, $x0$, n), which iterates $f(x)$, initialized by $x0$, n times, returns simplified – or better approximated – a vector with $n + 1$ components.

```
#1:   FIXED_POINT(2·COS(x), x, 1.5, 10)

#2:   [1.5, 0.1414744033, 1.980018354, -0.7957914253, 1.399439154, 0.3410396031, 1.884814923,
        -0.6177664155, 1.630348433, -0.1190338253, 1.985847670]
```

#3: $\text{FIXED\_POINT}\left(\frac{x + 2\cdot\cos(x)}{2}, x, 1.5, 10\right)$

```
#4:   [1.5, 0.8207372016, 1.092050620, 1.006691570, 1.038005274, 1.026942146, 1.030909002,
        1.029493846, 1.029999627, 1.029818978, 1.029883515]
```

Figure 2.1

The interested reader will see how to use the system of his choice and/or to program himself these different methods. Used between 1999 and 2011 at ÉTS, the TI-89 Titanium or Voyage 200 symbolic calculators from Texas Instruments were already showing (figure 2.2) how well we can do if we can store functions in memory and then recall them easily

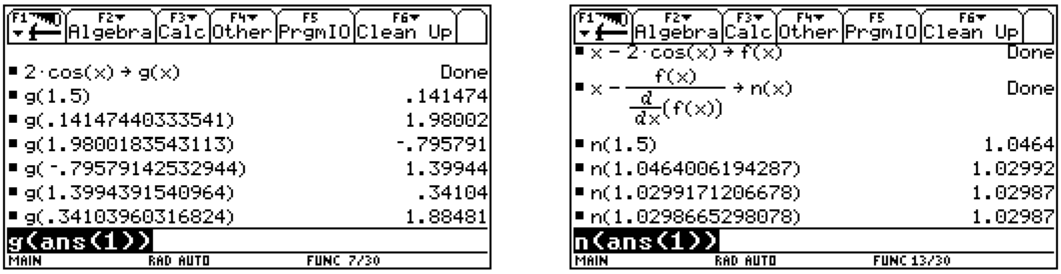


Figure 2.2

We can also visualize the fact that $g(x) = 2\cos x$ does not work: here with Nspire CX CAS.

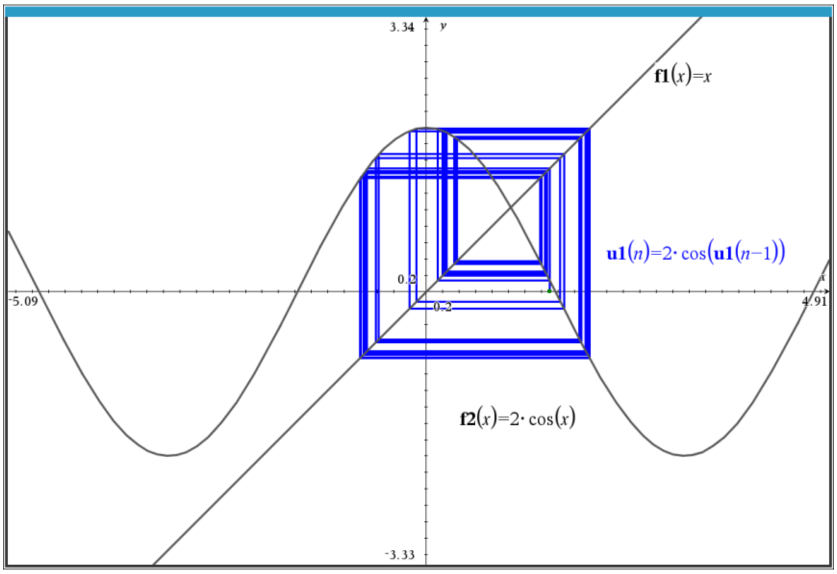


Figure 2.3

The functions "fixed_point" and "newton" are included in the kit_ets_mb library:

$$kit\_ets\_mb\backslash fixed\_point\left(\frac{x+2\cdot\cos(x)}{2},x,1.,10\right) \quad \begin{bmatrix} 1. \\ 1.0403 \\ 1.02611 \\ 1.0312 \\ 1.02939 \\ 1.03004 \\ 1.02981 \\ 1.02989 \\ 1.02986 \\ 1.02987 \\ 1.02987 \end{bmatrix}$$

Figure 2.4

2.3 Remark and reminder: Let's remember <u>Taylor's formula</u> for one variable functions: let f a differentiable function up to order $n+1$ in an open interval I containing the point a, then we have for any $x \in I$:

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x)$$

with the remainder (the error) $R_n(x)$ calculated as $R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$ for a c between x and a.

In fact, by posing

$$g(t) = f(x) - f(t) - f'(t)(x-t) - \frac{f''(t)}{2!}(x-t)^2 - \cdots \quad \frac{f^{(n)}(t)}{n!}(x-t)^n - R_n(x)\frac{(x-t)^{n+1}}{(x-a)^{n+1}}$$

it is easy to verify that g satisfies Rolle's theorem ($g(a) = g(x) = 0$). Therefore, there exists a number c between a and x such that $g'(c) = 0$. But $g'(t) = -\frac{f^{(n+1)}(t)}{n!}(x-t)^n + R_n(x)(n+1)\frac{(x-t)^n}{(x-a)^{n+1}}$, and it is sufficient to set $g'(c) = 0$ and solve it.

2.4 Theorem (**Newton's Method** and speed convergence) Let f be a function with a continuous second derivative on the interval I and with a zero r inside I, hence $f(r) = 0$. Suppose that there are positive numbers m and M such that $|f'(x)| \geq m$ et $|f''(x)| \leq M$ sur I.

If $x_1 \in I$ is close enough to r, say $|x_1 - r| < \frac{2m}{M}$, if

$$(1) \qquad x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \; n = 1,\ 2,\ \ldots$$

Then you have

$$|x_{n+1} - r| < \frac{M}{2m}(x_n - r)^2 \text{ et } \lim_{n\to\infty} x_n = r.$$

Proof: Taylor's formula implies that there exists a number c between x_n and r such that

$$f(r) = f(x_n) + f'(x_n)(r - x_n) + \frac{f''(c)}{2!}(r - x_n)^2$$

Because $f(r) = 0$, we get $|x_{n+1} - r| = \left| x_n - \frac{f(x_n)}{f'(x_n)} - r \right| = \left| \frac{f''(c)}{2f'(x_n)} \right| (x_n - r)^2 \leq \frac{M}{2m}(x_n - r)^2$.

Let's show, by recurrence on n, that we have $|x_n - r| \leq \frac{2m}{M}\left(\frac{M}{2m}|x_1 - r|\right)^{2^{n-1}} \to 0$ if $n \to \infty$. If $n = 1$, it is true, because $|x_1 - r| \leq \frac{2m}{M}\left(\frac{M}{2m}|x_1 - r|\right)^{2^0} = |x_1 - r|$. Assume that this is true for $n = k$, so assume that $|x_k - r| \leq \frac{2m}{M}\left(\frac{M}{2m}|x_1 - r|\right)^{2^{k-1}}$. Let's show that it is also true pour $n = k + 1$.

We have $\left|x_{k+1}-r\right| \le \frac{M}{2m}\left(x_k-r\right)^2 \le \frac{M}{2m}\left(\frac{2m}{M}\left(\frac{M}{2m}\left|x_1-r\right|\right)^{2^{k-1}}\right)^2 = \frac{2m}{M}\left(\frac{M}{2m}\left|x_1-r\right|\right)^{2^k}$ and the recurrence is completed. So, by choosing $\left|x_1-r\right| < \frac{2m}{M}$, we have $\lim_{n\to\infty}\left|x_n-r\right| = 0$, which ends the proof. ♦

2.5 Examples: (details in class) : since Newton's method consists in iterating the function

$$x-\frac{f(x)}{f'(x)},$$

We find this method implemented in *Derive*. Example from above gives after approximating the function NEWTON(*f*, *x*, *x0*, *n*) for *f*(*x*) = 0.

```
#5:   NEWTON(x - 2·COS(x), x, 1.5, 10)

#6:   [1.5, 1.046400619, 1.029917120, 1.029866529, 1.029866529, 1.029866529,

        1.029866529, 1.029866529, 1.029866529, 1.029866529, 1.029866529]
```

Figure 2.5

We have “exported“ this function to TI-NspireCAS:

```
kit_ets_mb\newton(x−2· cos(x),x,1.5,10)        [1.5
                                                1.0464
                                                1.02992
                                                1.02987
                                                1.02987
                                                1.02987
                                                1.02987
                                                1.02987
                                                1.02987
                                                1.02987
                                                1.02987]
```

Figure 2.6

Fixed point method and Newton's method as well are extended to systems of equations and computer software provides programmed functions to execute them. (obviously, the proofs of these results involve norms of Jacobian matrices). Thus, if we try to solve a system of equations (non-linear and very often non-polynomial system) given by

$$\begin{cases} f_1\left(x_1, x_2, \ldots\right) \\ f_2\left(x_1, x_2, \ldots\right) \\ \vdots \\ f_n\left(x_1, x_2, \ldots\right) \end{cases}$$

and if we want to apply Newton's method, we have to iterate the vector where $\mathbf{x}-\mathbf{J}^{-1}\mathbf{F}$ where

$$\mathbf{x} = \left[x_1 \;\; x_2 \;\; x_3 \ldots \quad \left[\ \right]_{i,j=1}^{n} \quad \ldots \quad .\right.$$

xo is the initial point. Matrix **J** is the Jacobi matrix of the system. Obviously, an implementation will replace the calculation of $\mathbf{J}^{-1}\mathbf{F}$ by a Gauss-Jordan reduction of the augmented matrix from which the last column will be extracted having taken care to insert the point prior to the reduction! We will illustrate in class with the following systems:

2.5.1 $\{x^2+y+z-3=0, \quad x+y^2+z-3=0, \quad x+y+z^2-3=0$

Note that this system is analytically solvable and has eight real solutions of which two are easy to find by inspection. It is particularly interesting to make a connection here with calculating with more variables. Both parabolic cylinders $z=3-x^2-y,\ z=3-x-y^2$ intersect in a curve which consists of two parabolas - which can be checked - $r_1(t)=\left[t,\ t, 3-t^2-t\right]$ and $r_2(t)=\left[t, -t+1,\ 2-t^2+t\right]$ in parameter form. By substituting into the equation of the third surface, we obtain the equation $t^4+2t^3-5t^2-4t+6=0$ which has solutions $\pm\sqrt{2}$, 1 and -3. This provides the eight solutions! Here is an illustration of the two parabolas coming from the intersection of the first two cylinders:

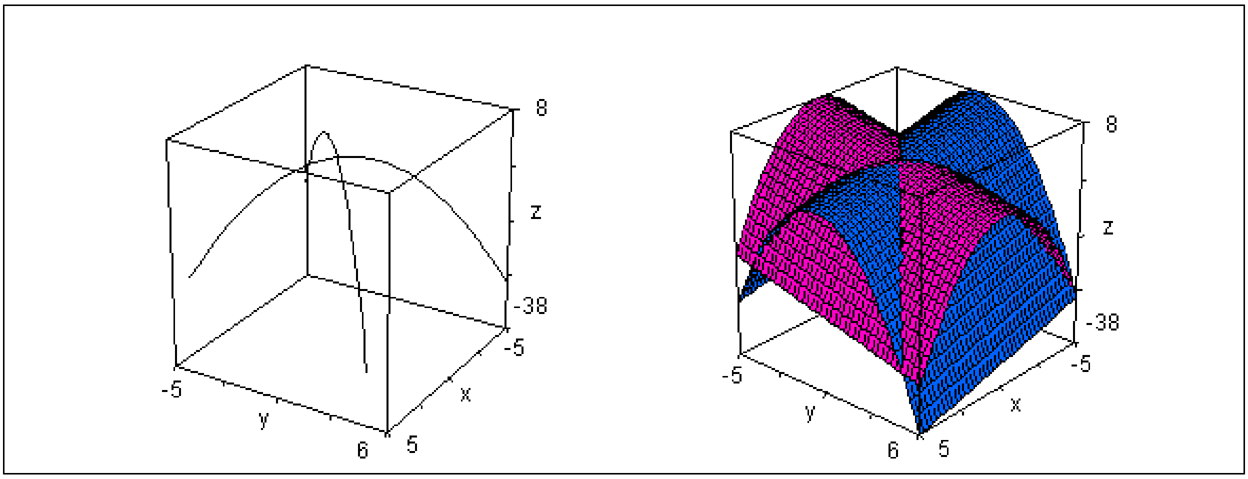


Figure 2.7

Here are the two parabolas "hitting" the third cylinder, namely (each of the two parabolas meets the cylinder four times):

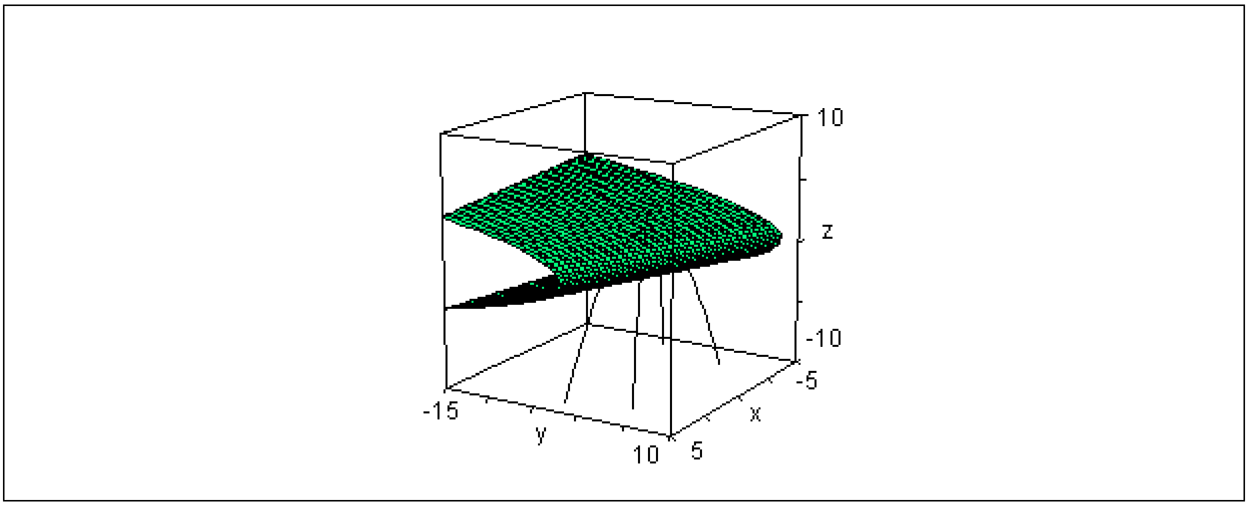


Figure 2.8

The fact that this polynomial system has eight solutions can be demonstrated mathematically: computeralgebra systems are equipped with "Gröbner base" type functions, which reduce a polynomial system into another equivalent one. Its specialty is to contain an equation with only one unknown. We can therefore apply the fundamental theorem of algebra and then substitute in the other equations. This allows us to know the exact number of solutions. In the example we are interested in, let's see what *Derive 6* will give. Its "SOLUTION" function gives the eight solutions in matrix form:

#1: $\text{syst} := \left[x^2 + y + z - 3,\ x + y^2 + z - 3,\ x + y + z^2 - 3\right]$

#2: $\text{SOLUTIONS}(\text{syst}, [x, y, z]) = \begin{bmatrix} 1 & 1 & 1 \\ -3 & -3 & -3 \\ \sqrt{2} & \sqrt{2} & 1 - \sqrt{2} \\ \sqrt{2} & 1 - \sqrt{2} & \sqrt{2} \\ \sqrt{2} + 1 & -\sqrt{2} & -\sqrt{2} \\ -\sqrt{2} & \sqrt{2} + 1 & -\sqrt{2} \\ -\sqrt{2} & -\sqrt{2} & \sqrt{2} + 1 \\ 1 - \sqrt{2} & \sqrt{2} & \sqrt{2} \end{bmatrix}$

Figure 2.9

And its "GROEBNER_BASIS" function allows to see that there are exactly six different values of *z* (but by substituting each of them, eight values are generated for *x* and *y*):

#3: $\text{GROEBNER\_BASIS}(\text{syst}, [x, y, z])$

#4: $\left[z^6 - 10\cdot z^4 + 4\cdot z^3 + 19\cdot z^2 - 8\cdot z - 6,\ y\cdot(2\cdot z^2 - 4) + z^4 - 5\cdot z^2 + 6,\ y^2 - y - z^2 + z,\ x + y + z^2 - 3\right]$

#5: $\text{SOLUTIONS}(z^6 - 10\cdot z^4 + 4\cdot z^3 + 19\cdot z^2 - 8\cdot z - 6, z)$

#6: $[1, -3, \sqrt{2}, \sqrt{2} + 1, -\sqrt{2}, 1 - \sqrt{2}]$

Figure 2.10

2.5.2 $\begin{cases} x\cos 2 + 2\sin y = \dfrac{1}{2} \\ x\,y = 2 \end{cases}$

The previous system is reducible to a single equation with one unknown but it must necessarily be solved numerically and has an infinite number of solutions as the reader will be able to verify.

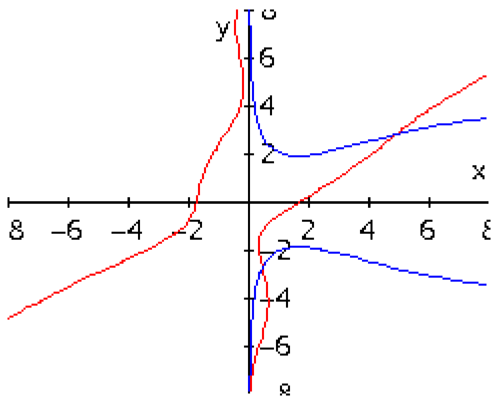
2.5.3 Here follows a *Derive* session for the system

$$\begin{cases} 3x\,y - 2x^2 + 4\sin(y) + 6 = 0 \\ 3x^2 - 2x\,y^2 + 3\cos(x) + 4 = 0 \end{cases}$$

Here, the function executing Newton's method requires vector notation: in *Derive* one has to enter NEWTONS(**f**, **x**, **x0**, *n*), where **f** is the vector of expressions to be set to zero, **x** is the vector of the unknowns, **x0** is the initial vector and *n* is the number of iterations to be performed. To get an idea of the starting point, we draw, in the same window, each of the two curves defined by the above equations: Curve #8 is the red one. This is analytically predictable since it is impossible that $x = 0$ in this equation...!

#8: $3\cdot x\cdot y - 2\cdot x^2 + 4\cdot \mathrm{SIN}(y) + 6 = 0$

#9: $3\cdot x^2 - 2\cdot x\cdot y^2 + 3\cdot \mathrm{COS}(x) + 4 = 0$



#10: $\mathrm{NEWTONS}\left(\left[3\cdot x\cdot y - 2\cdot x^2 + 4\cdot \mathrm{SIN}(y) + 6,\ 3\cdot x^2 - 2\cdot x\cdot y^2 + 3\cdot \mathrm{COS}(x) + 4\right], [x, y], [5, 3], 4\right)$

#11: $\begin{bmatrix} 5 & 3 \\ 4.963335864 & 2.821759042 \\ 4.954283759 & 2.812083609 \\ 4.954246603 & 2.812041878 \\ 4.954246603 & 2.812041878 \end{bmatrix}$

Figure 2.10

We can also do this example with Nspire and still produce the graph of the curves since implicit 2D graphics are not really required.

Indeed, it is possible to solve for x in the first equation, which gives two branches, and do the same for y in the second equation, also giving two branches (see comment to figure 2.12). Figure 2.11 presents the work and use of the Newtons2 function of the kit_est_mb library. Note that the Nspire solver, without a starting point, finds the solution located in the fourth quadrant. And by giving it an "adequate" starting point, the solver finds the solution located in the first quadrant.

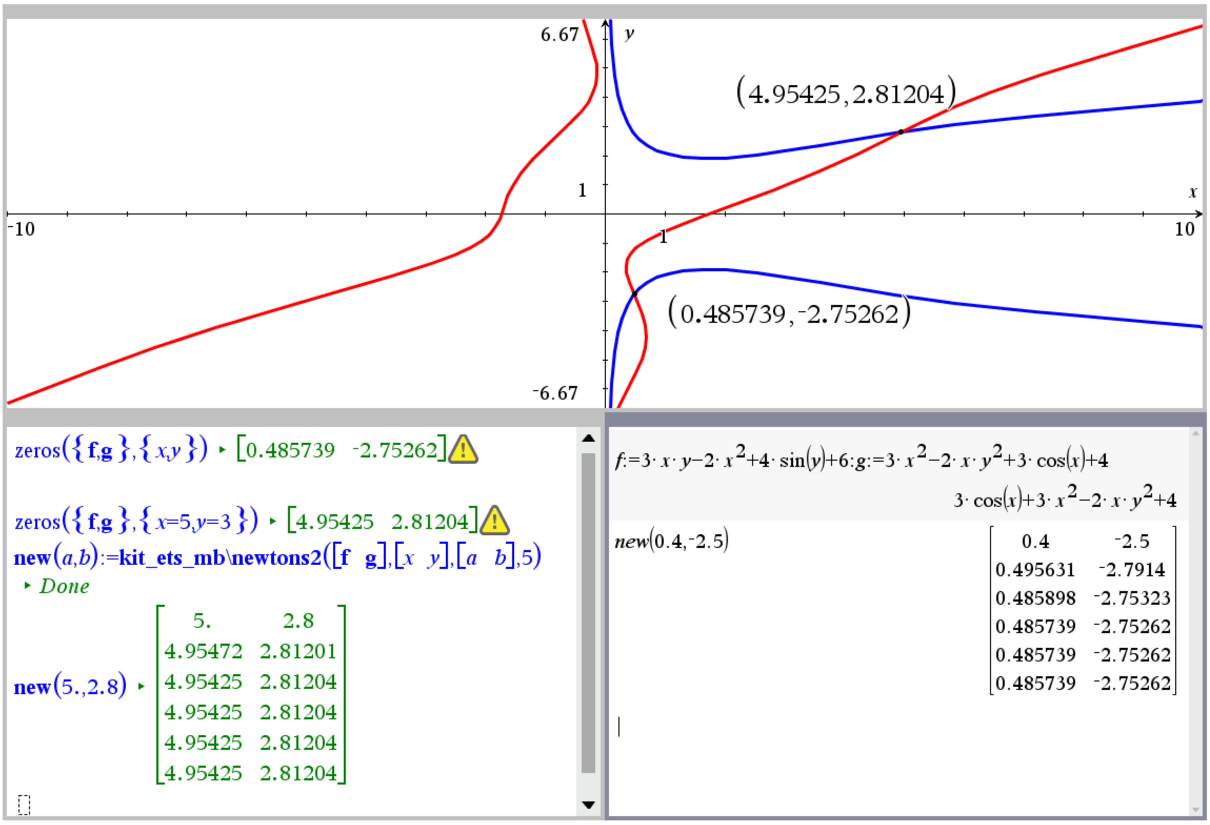


Figure 2.11

Comment: The red curves must be entered as relations!

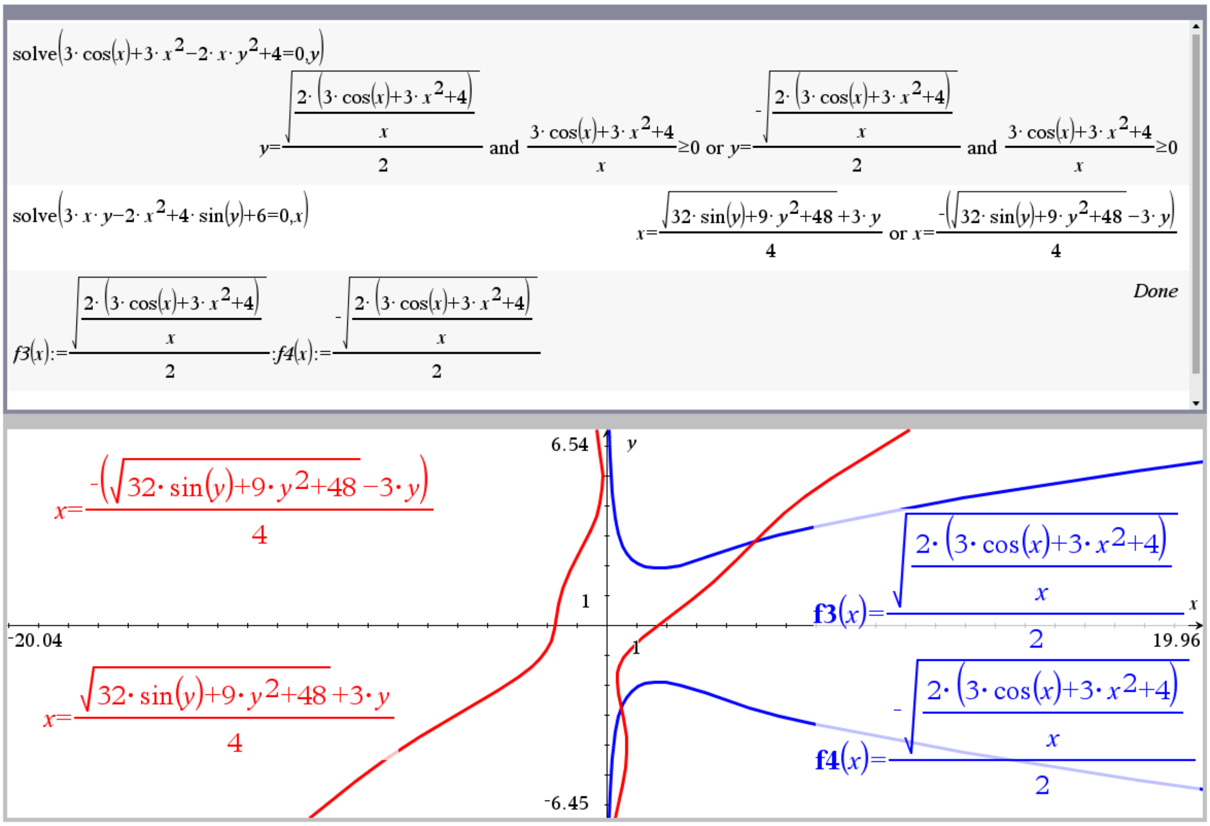


Figure 2.12