# Solving equations and systems of equations: Newton's method

Valery Ochkov, Yulia Chudova

National Research University "MPEI"

*I <…> was solving some long algebraic equation on a black board. In one hand I held Franker's torn soft "Algebra", in the other*—*a small piece of chalk, with which I had already soiled both hands, face and elbows ...*

*Leo Tolstoy "Youth"*

The main reproach cast by society against school is its isolation from life. In school they teach one thing, but in life you have to deal with something completely different.

With regard to school mathematics, then this reproach may be presented as follows. At school, they teach analytical methods for solving far-fetched equations and their systems, but in life one has to apply numerical (approximate) methods to real equations, which are not even mentioned at school.

Here is one example from classic cinema. In the cult Italian film "Amarcord" (1973), a teacher explains to a student (see Figure 1) how to solve the equation by converting it to the form a x2 + b x + c = 0. You need to use the numerical values of the coefficients a, b and c and substitute them into the well-known formula for the roots of a quadratic equation with a discriminant under the square root.

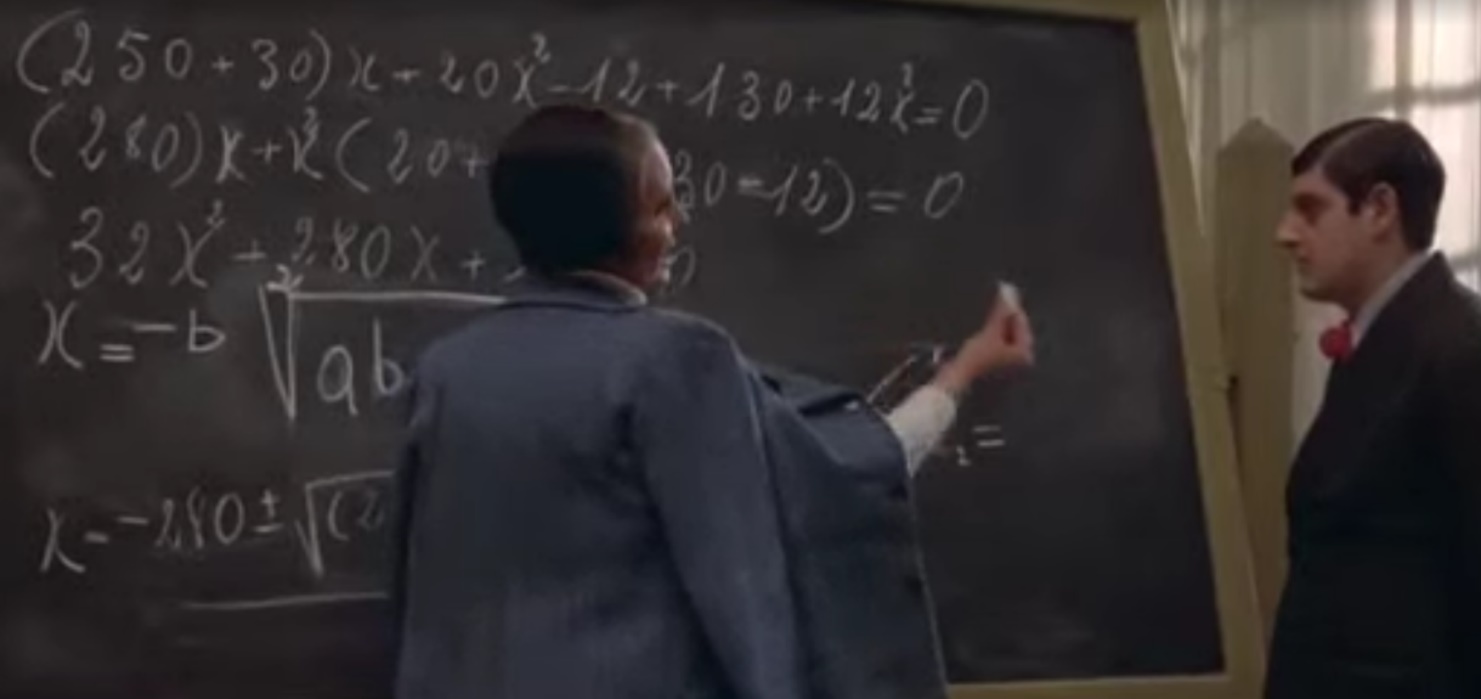


Fig. 1. Frame from the film "Amarcord"

In the past, numerical methods for solving equations were not considered at school because they were very laborious even when working with electronic calculators. But now, computers with modern mathematical programs have become available in schools, and they can be used not only in computer science lessons, but also in mathematics lessons for the numerical solution of equations and their systems. Moreover, the teacher can now write on an electronic board with a special pen, rather than with chalk.

For now, we only touch on algebraic equations. But what will be said in the article also applies to differential equations [1].

Let's install the domestic freely distributed mathematical program SMath (www.smath.com) on a computer in a couple of minutes and use it to show how you can quickly and clearly numerically find the zero of a function (find its root) or solve a system of equations using Newton's method.

Figure 2 shows the essence of Newton's method. The analyzed function y(x) is set – here a cubic polynomial[[1]](#footnote-1)—and the first assumption (x1 = -2) is the starting point on the graph through which the tangent is drawn. An alternative name for Newton's method is the method of tangents. Isaac Newton (1642-1727) together with Gottfried Leibniz (1646-1716) created differential calculus[[2]](#footnote-2) with its derivatives and tangents! Our tangent crosses the x-axis at the point x2 = –0.9844, which will be considered the second approximation. A new tangent is drawn through this point, which intersects the x-axis at the point –0.1767. These actions are repeated until the value of the function at the next point becomes equal to zero (approximately zero!), when our two points merge into one and both are on the x-axis. In our calculation, this required 16 iterations, six of which (n=1, 2, 3, 12, 14, and 16) are shown in Figure 2. At the 15th iteration, two separate points and two separate values of their abscissas are still visible—red and black[[3]](#footnote-3). But on the 16th, last iteration, the points themselves visually coincide, as do the two values of their abscissas, 5.1685,—red is written over black. This is the desired zero of the function.

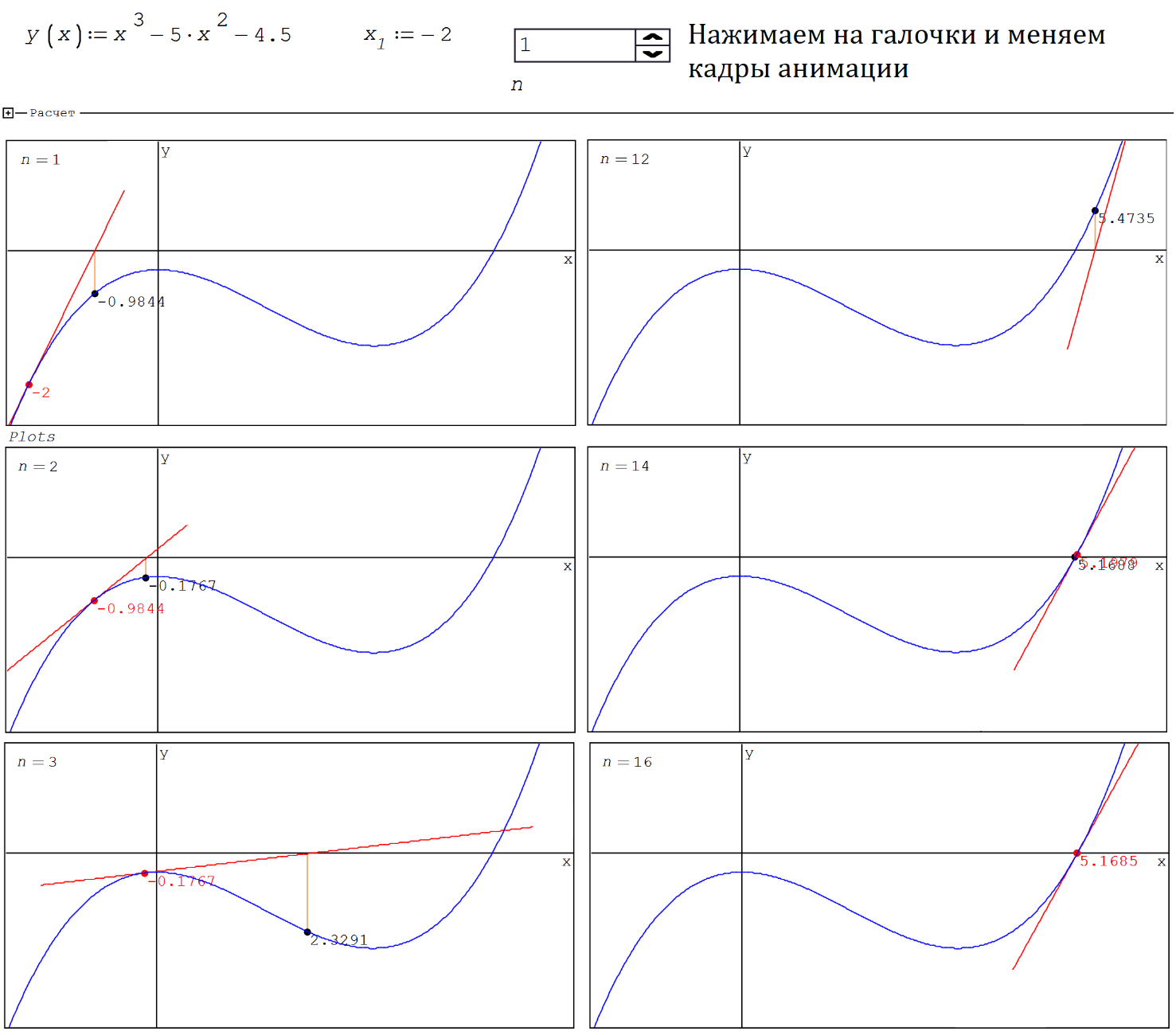


Fig. 2. Graphical display of the function approaching zero

In figure 2 the function y(x) and its first approximation to zero x1 are given by the := (assign) operator. And the iteration number n is set through the Controls/Numeric up-down control, which is more convenient for our task. If you click on the checkboxes of this interface element (Controls)—normal and inverted, then the value of the variable n will instantly change down (minus 1) or up (plus 1) with a single step[[4]](#footnote-4). This will allow us to see our lines illustrating Newton's method, almost like in animation. But you can also create a real animation, save it as a gif-file, for example, and show it separately without linking to the SMath package, which is described in chapter 1 of the tutorial [2].

Figure 3 shows the operators hidden in Figure 2 in the collapsed area named Calculation.

There, firstly, the derivative of the analyzed function is obtained analytically, then the equation of the tangent at the point x1 is introduced, and then a function called Zero (zero) is created with an infinite loop, the exit from which (break) is determined by the user by setting the value of the variable n—the number of iterations (see Figure 2).

* The Plots variable stores data for plotting seven graphs:
* graphics of the function itself;
* graphics of its tangent—a straight line;
* a red dot with a size of 15 units, for the previous approximation;
* a black dot with a size of 15 units, for the current approximation;
* a vertical line connecting the y-axis with the point of the current approximation (data for the graph in the form of a square matrix);
* red inscription (with font size 10) of the value of the argument at the point of the previous approximation;
* black inscription (with font size 10) of the value of the argument at the point of the current approximation.

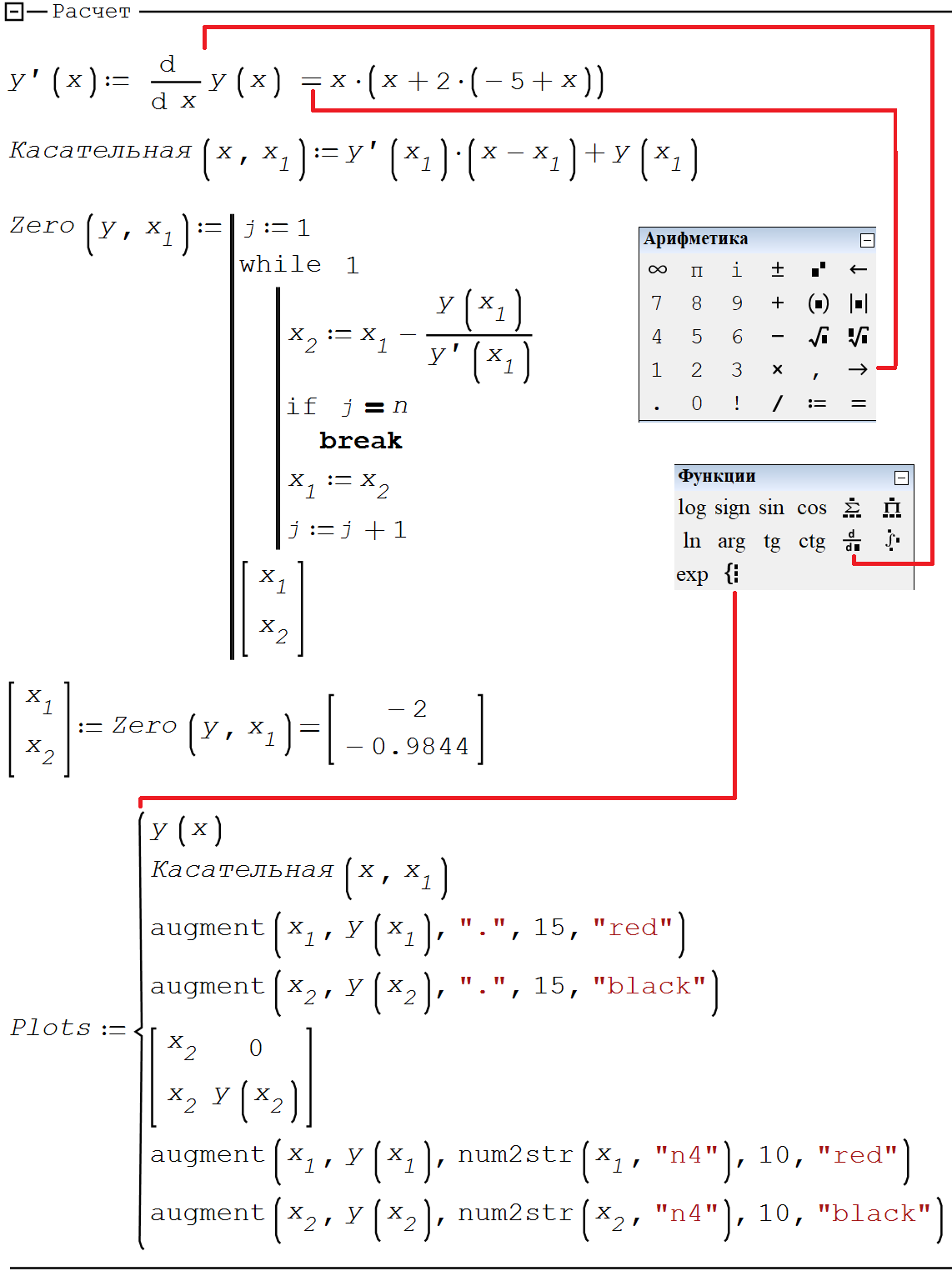


Fig. 3. The program for finding the zero of a function by Newton's method and operators for plotting graphs

The tangent may be replaced by a secant by specifying not one, but two points for the initial guess. If these points are located close to each other and both are to the left or to the right of the desired zero, then the secant method (or the chord method—that is its name) will not differ much in appearance from the tangent method. But if the desired zero is placed between the points of the initial assumption, then the pattern of searching for the zero of the function changes dramatically—see Figure 4. It will be somewhat reminiscent of another widely used method of numerically searching for the zero of a function—the bisection method. Its description is beyond the scope of this article. But an inquisitive reader can find the necessary information on the Internet.

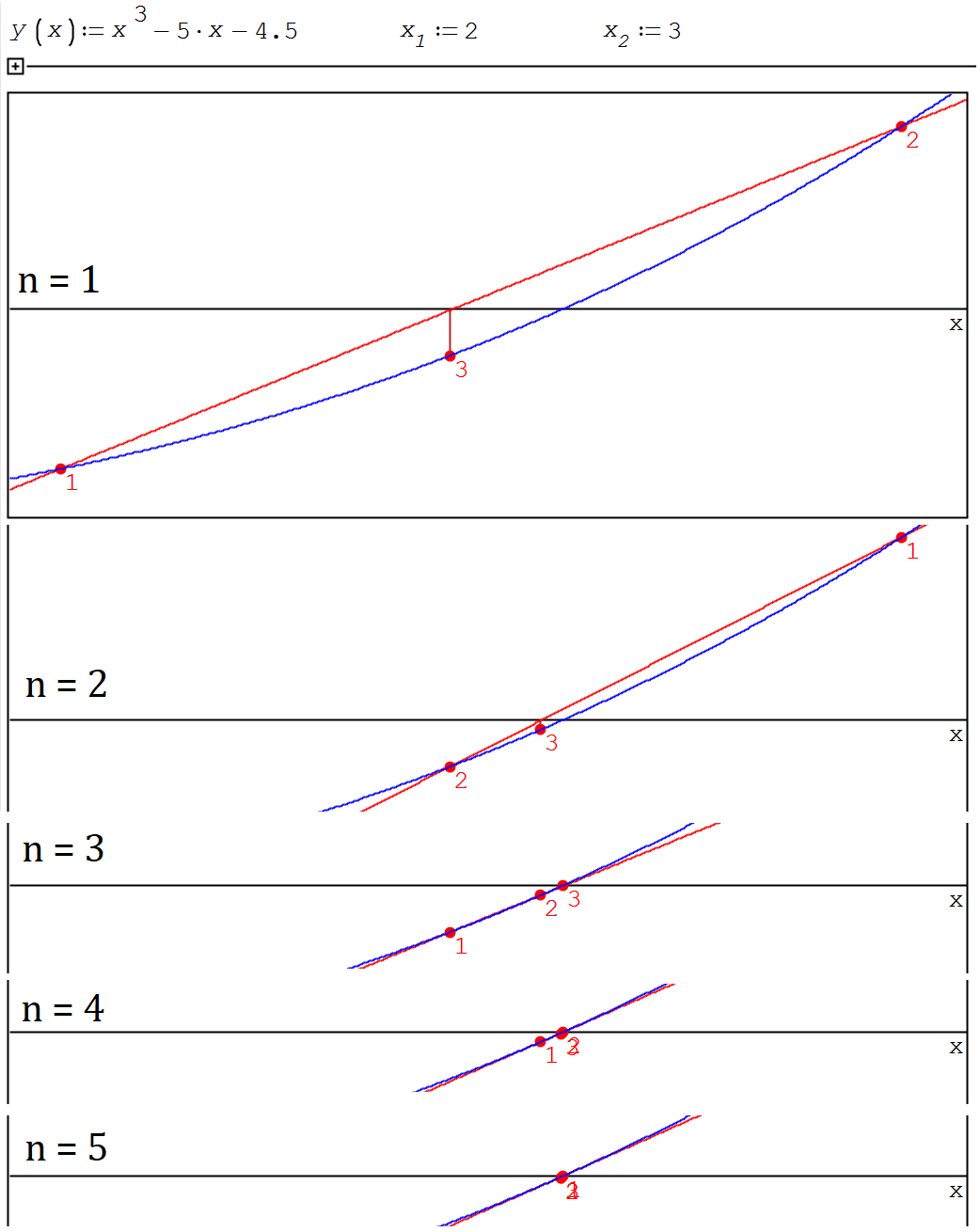


Fig. 4. Graphical illustration of the secant method

Newton's method for a function of one argument is well described on the Internet and we have not come up with anything new here. We only implemented it for the first time in the SMath environment.

In here, for functions of two or more arguments (for systems of equations), there is almost no information on the Internet. Let's fill this gap.

Figure 5 shows how two equations of Cassini ovals (f) are introduced into the calculation— Leo Tolstoy's oval [3, 4] (f1) and Bernoulli's lemniscate (f2). These two closed fourth-order curves are shown in Figure 5. Next, a vector function named F and a square matrix of its partial derivatives with respect to two arguments are specified. This matrix is called the Jacobian (J) matrix and is calculated using the built-in Jacobian function in SMath. Then the first guess is set—the first elements of the vectors X and Y. The remaining elements of these vectors are set in the for loop in successive iterations.

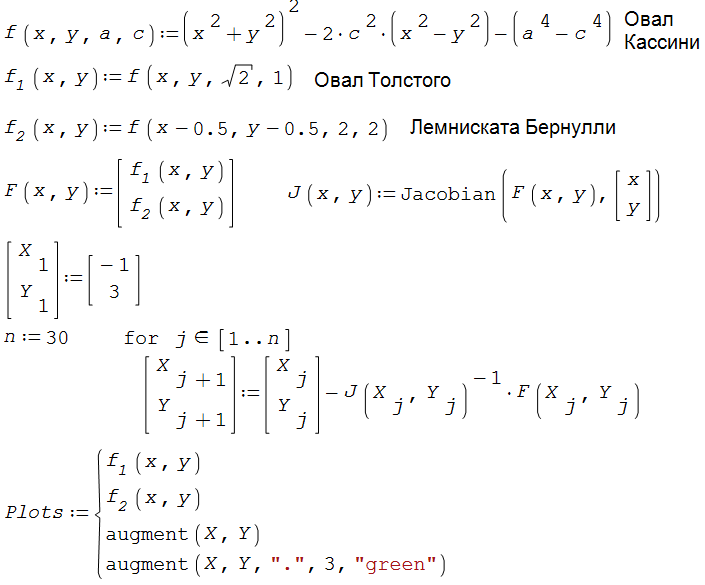


Fig. 5. Steps of searching for four roots (solutions) of the system of equations

When implementing Newton's method for a system of two equations, it is necessary to work not with a single derivative, but with a set of derivatives, known as the Jacobian matrix. This 2 by 2 square matrix stores the partial derivatives of our two analyzed functions with respect to two arguments. Everything else is a matter of technology (computer technology). In Figure 6 the green traces show the numerical searches for solutions from different starting points.

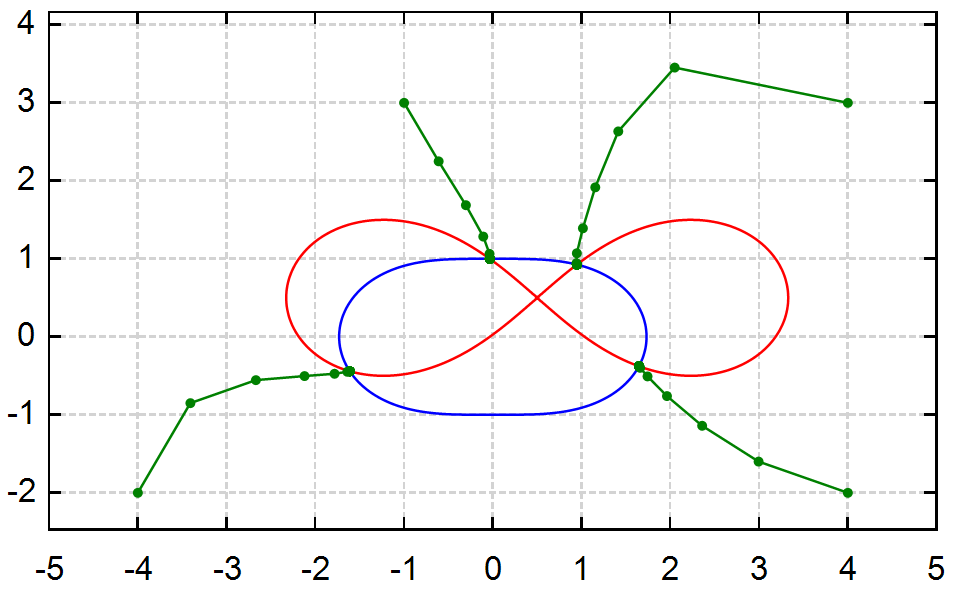


Fig. 6. Graphical display of the solution of a system of two equations by Newton's method

If there were three equations, then the Jacobian matrix would consist of three rows and three columns. In this case, it would be necessary to graphically display the solution not with two curves, but with three surfaces on a 3D graph, the mutual intersection of which would mark the solution.

The sequential iterative construction in Figure 5 repeats the entry in Figure 3 above the if statement, where a fraction is subtracted from the previous approximation, the numerator of which is the analyzed function, and the denominator is its derivative. When working with two equations, division is replaced by matrix multiplication of the inverse Jacobian matrix (a matrix raised to the negative first power) by the function F, a vector that stores the analyzed functions. If a number is multiplied by its reciprocal, the result is one. If a matrix is multiplied by its inverse matrix, then we get an identity matrix, the main diagonal of which stores ones, and the rest is zeros.

The area shown in Figure 6 can be colored in four colors, marking the zones of “attraction” of points of four solutions [5]. But here a fifth area may also appear—one from which no solution will be found. This is, for example, the origin, where the calculation will be interrupted by an error message.

Figure 7 shows the zones of attraction (zones of influence) of solutions to a system of two equations, the first of which is the "equation of the heart" (sixth order curve), and the second is the equation of an arrow penetrating the heart. It results in quite an interesting portrait in three colors: light purple is the zone of attraction of the left root, dark purple is the right root, and black is the zone of no solution. Yellow circles at the temples of the portrait are two solutions.

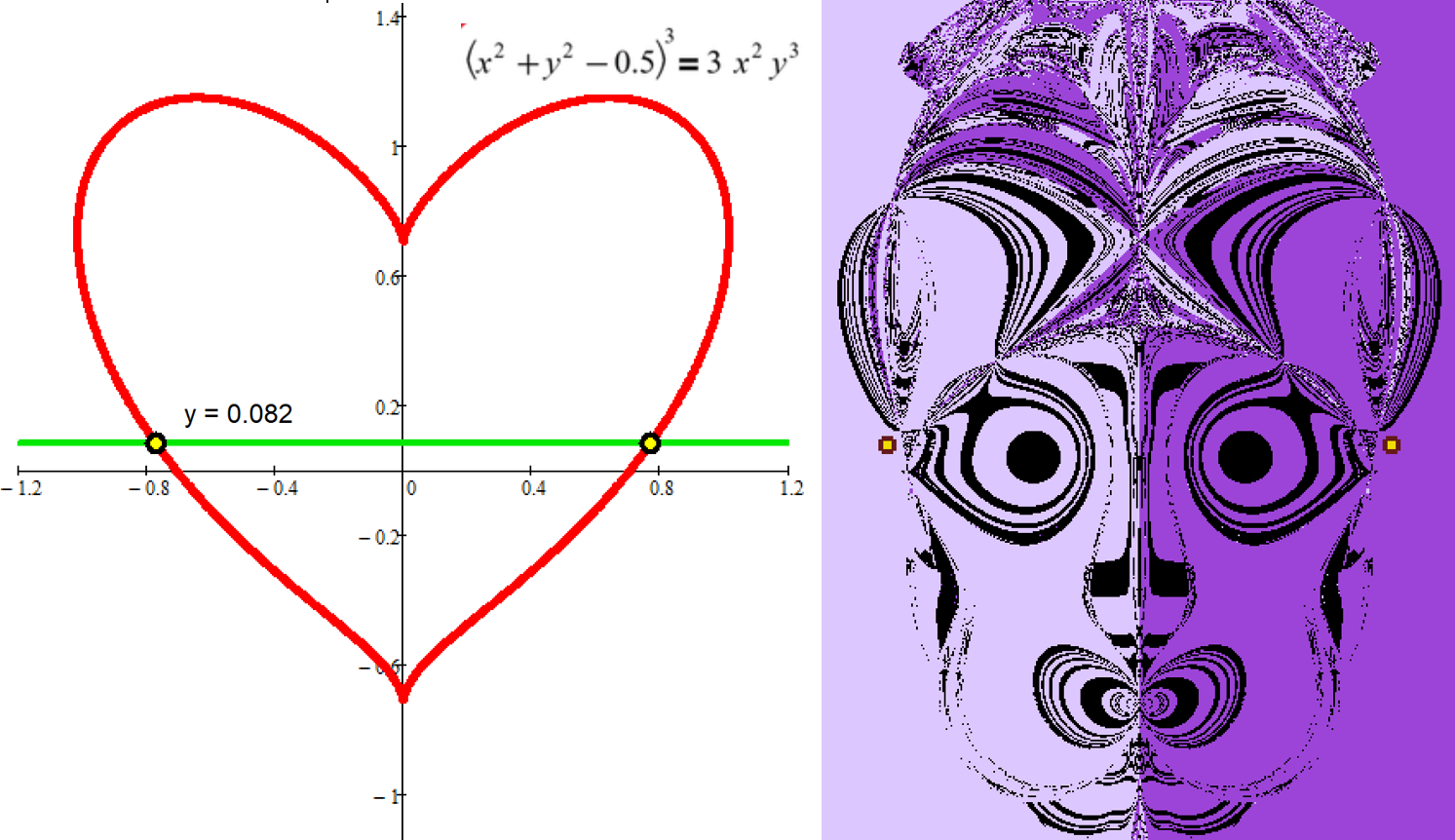


Fig. 7. Mystical portrait of the solution of a system of two equations

Figure 8 shows the analytical (symbolic) solution of our cubic equation using the site wolframalpha.com, which is very popular with students and schoolchildren. The number 4.5 is replaced by the constant c. The site gave three solutions, one of which is shown in the figures, and the other two (complex expressions) are not shown.

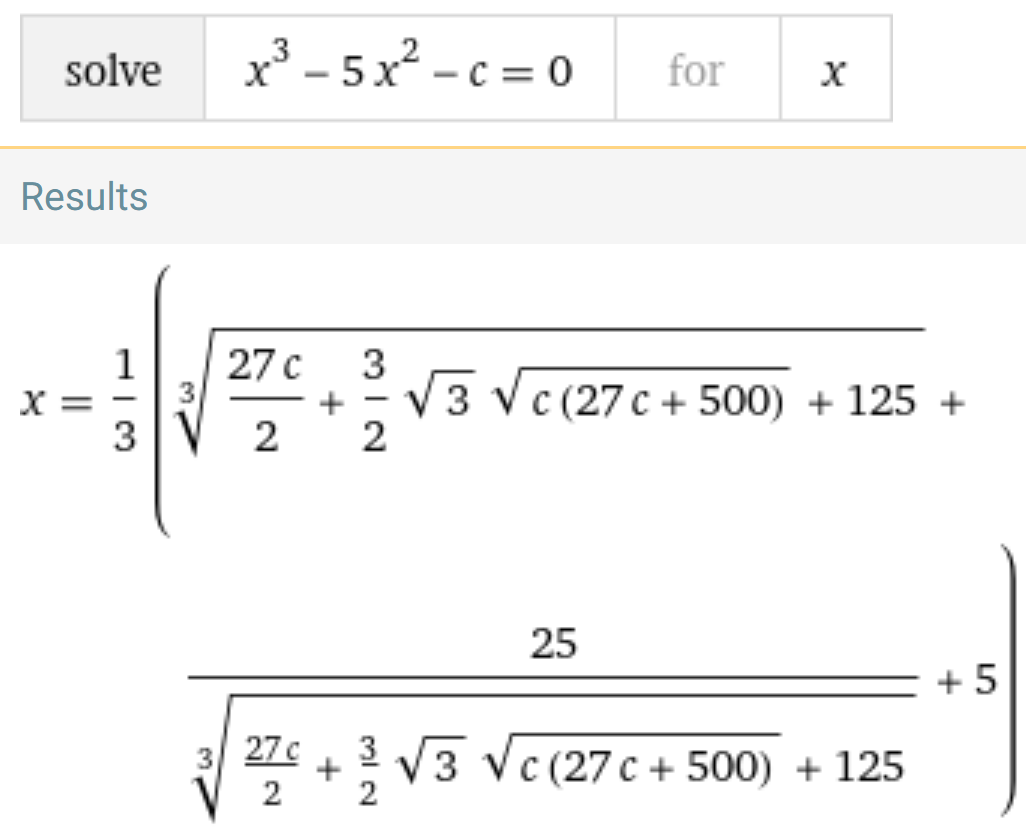


Fig. 8. Analytical solution of the equation

The best way to solve the problem is a combination of symbolic, numerical and graphical approaches, which, by the way, is what we did. So in Figure 3 the derivative of the function was calculated analytically (symbolically), and the rest was done numerically. Graphs also helped us solve problems—to localize zeros and roots. Such a hybrid solution of the problem [6] is the way to success!

At present, many complex algorithms for solving equations and their systems have been developed. Now it is not clear why this was done—to improve the accuracy of the calculation and / or to reduce the area of the black zones shown in Figure 7, or simply—to increase the speed of work on old low-speed computers. Modern high-speed computers have caused a renaissance of the good old simple methods of solving problems. Newton's method, for example. It is almost impossible for the uninitiated to understand complex algorithms, and simple algorithms are visible at a glance, which is very important for educational purposes.

Sophisticated modern methods can be likened to modern jet liners. But I really want to make the simplest wdesign like a hang glider and fly on it—which is what we have tried to do.

Literature:

1. V.F. Ochkov and E.P. This is a terrible word difura... // Informatics at school. No. 1. 2015. P. 55-58 (http://www.twt.mpei.ac.ru/ochkov/ODE.pdf)
2. Points V.F. and others. Information technology engineering calculations: SMath & Python. Lan Publishing. 2023 (http://twt.mpei.ac.ru/ochkov/EC-SMath.pdf)
3. Leo Tolstoy and mathematics / V. F. Ochkov, N. A. Ochkova. Moscow: MPGU, 2023.—208 p. (http://www.twt.mpei.ac.ru/ochkov/Tolstoy-Math-3.pdf)
4. https://mathcurve.com/courbes2d/cassini/cassini.shtml
5. V. F. Ochkov, Yu. V. Chudova, and N. R. Umirova. Portrait of the roots of the system of equations // Mathematical Education No. 3 (103), 2022. P. 33-46 (http://www.twt.mpei.ac.ru/ochkov/Portrait-Roots.pdf)
6. V. F. Ochkov, A. V. Bobryakov, and S. N. Khorkov, Acoust. Hybrid problem solving on a computer // Cloud of Science. Volume 4 No. 2. 2017. P. 5-26 (http://twt.mpei.ac.ru/ochkov/Hybrid.pdf)

1. Here, a cubic rather than a square polynomial (as in Figure 1) is taken with a certain hump, through which our tangent will “roll” several times until it gets to the solution. [↑](#footnote-ref-1)
2. https://en.wikipedia.org/wiki/Leibniz–Newton\_calculus\_controversy [↑](#footnote-ref-2)
3. In the novel by Stendhal, "Le Rouge et le Noir", you can see some iterations of the protagonist between two career lines—military (black) and church (red). [↑](#footnote-ref-3)
4. This setting becomes available by pressing the right mouse button. The Numeric up-down control becomes available after loading the corresponding plugin. [↑](#footnote-ref-4)