Annales Academiae Paedagogicae Cracoviensis
Folia 16 Studia Mathematica III (2003)

Daniela Bittnerová

The role of notation in mathematics

Abstract. In this paper, the historical development of the Cardan's rule notation is presented. Some descriptions with respect to the linear and/or structured form of texts are compared, and also the exactness of these descriptions is discussed. It is interesting to look at the way of finding the solution. An English transaltion of the first general method for solving cubic equations in Cardan's Ars Magna is quoted and compared with its contemporary presentation. This comparison could be useful for presentation of a new mathematical problem to students or for writing a mathematical text for students.

If we deal with any (mathematical) textbook we are interested not only in its content but we also look at the formal aspect. At the first sight some books are nice and interesting, other books are uninteresting.

The most frequent form of a mathematical text is the linear or almost linear one. Then the reader usually modifies this text according to personal needs. The reader studies the text with a pen in hand and proceeds along the lines, crossing the lines, in blocks etc. It is true that mathematicians do not make the situation easy for the reader at this point.

Let us look at one algebraic theme and compare several different ways of its presentations.

When the Italian mathematician Hieronim Cardan published the first general method for solving cubic equations in his Ars Magna (1545), he described the solution of the equation

$$x^3 + px = q, \quad p \in \mathbb{R}, \quad q \in \mathbb{R},$$

as follows (translation from Latin):

> Raise the third part of the coefficient of the unknown to the cube, to which you add the square of half the coefficient of the equation, & take the root of the sum, namely the square one, and this you will copy, and to one [copy] you add the half of the coefficient that you have just multiplied by itself, from another [copy] you subtract the same half, and you will have the Binomium with its Apotome, next, when the cube root

Supported by the project MSM 245100302 "Discrete mathematics and its application".

of the Apotome is subtracted from the cube root of its Binomium, the remainder that is left from this, is the determined value of the unknown.

This description is strikingly clumsy, but at that time no better method of solving cubic equations was available. For our use, it is not essential that the text above is not original but an English translation from Latin. Besides "clumsiness" of this text, we can find some doubtful wordings there (for example, "the half of the coefficient" itself, or its square, has to be added to and subtracted from the copies"). The Cardan's wording in natural language is original and romantic, but it could occure utter gibberish to a mathematically uneducated reader, in comparison with the professional mathematical text. Furthermore, the linearity and complexity of the whole sentence makes the rule unintelligible and difficult to read.

Let us compare the quoted description above with the same rule in the present-day notation: ([4], 298_{11-9})

SOLUTION OF THE EQUATION $x^3 + px = q$.
Let $c = \sqrt{d}$, where $d = \left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2$, and let $b = c + \frac{q}{2}$ and $a = c - \frac{q}{2}$.
Then $x = \sqrt[3]{b} - \sqrt[3]{a}$ is a root of the equation.

Obviously, conciseness is the advantage of this notation. Also, the algebraic description using parentheses is quite unambiguous and makes possible to manipulate the formula for x. Thus:

Raising to power the expression $x = \sqrt[3]{b} - \sqrt[3]{a}$ we obtain

$$\begin{aligned} x^3 &= b - 3\sqrt[3]{b^2a} + 3\sqrt[3]{ba^2} - a = b - a - 3\sqrt[3]{ab}\left(\sqrt[3]{b} - \sqrt[3]{a}\right) \\ &= \lfloor\text{substituting } x \text{ and } q\rceil = q - 3\sqrt[3]{ab}x, \end{aligned}$$

because

$$q = 2(b - c) = 2(c - a). \text{ From here it is } q = b - a.$$

As

$$ba = \left(c + \frac{q}{2}\right)\left(c - \frac{q}{2}\right) = c^2 - \left(\frac{q}{2}\right) = \lfloor\text{according to the premise}\rceil = \left(\frac{p}{3}\right)^3,$$

we see that $x^3 + px = q$ really.

To verify this rule high-school mathematics was sufficient; a similar verification is impossible for the formulation in natural language.

A high-school textbook published in 1952 ([5], 509_{15}-516) investigates the conditions of existence of solutions of the cubic equation $x^3 + px + q = 0$ in detail. These conditions and types of roots of the equation are determined by means of local extremes of the function $f(x) = x^3 + px + q$. It is interesting that respecting the need to modify and ability to read a mathematical text, the commentaries are written linearly, the important mathematical expressions are mentioned on separately in the left or right-hand margin (however, quite inconsistently). The results of the investigation are presented in linear form (so rather not easy to survey), but at the end they are summarised in a well-arranged table accompanying the investigated graph illustrating the situation.

Let us recall the relevant part, i. e. **find the roots** of the equation $x^3 + px + q = 0$:

If p and q are real numbers the equation $x^3+px+q = 0$ has either three real roots or one real and two complex conjugated. Let us differentiate the function $f(x) = x^3 + px + q$. Then $f'(x) = 3x^2 + p$.

For $p > 0$, it is $f'(x) > 0$ and the function f is increasing for each real number x. Thus the function has only one real root (the intersection with the x-axis) because passing through all points from $-\infty$ to $+\infty$ the function changes the sign from minus to plus.

For $p < 0$, the function has the local $\begin{bmatrix} \text{maximum at the point} \\ \text{minimum} \end{bmatrix} \begin{bmatrix} x = -\sqrt{-\frac{p}{3}}, \\ x = \sqrt{-\frac{p}{3}}. \end{bmatrix}$

Putting these values for x in function f we obtain

$$f_{max} = q - \frac{2p}{3}\sqrt{-\frac{p}{3}} \quad \text{or} \quad f_{min} = q + \frac{2p}{3}\sqrt{-\frac{p}{3}}.$$

Let us simplify the product

$$f_{max} f_{min} = \left(q - \frac{2p}{3}\sqrt{-\frac{p}{3}}\right)\left(q + \frac{2p}{3}\sqrt{-\frac{p}{3}}\right) = q^2 + \frac{4p^3}{27}.$$

If the values f_{max}, f_{min} have the same signs, then

$$q^2 + \frac{4p^3}{27} > 0, \quad \text{i.e.} \quad \frac{q^2}{4} + \frac{p^3}{27} > 0,$$

and the equation has only one real root

$$x \in \left(-\infty; -\sqrt{-\frac{p}{3}}\right) \cup \left(\sqrt{-\frac{p}{3}}; +\infty\right).$$

If $f_{max} > 0$ and $f_{min} < 0$, it is $\frac{q^2}{4} + \frac{p^3}{27} < 0$. Then the signs of the function in the end or extremal points are:

$f(-\infty)$	f_{max}	f_{min}	$f(+\infty)$
$-$	$+$	$-$	$+$

The given equation has evidently three real roots.

If $\frac{q^2}{4} + \frac{p^3}{27} = 0$ the investigated equation has one double-root $x = \pm\sqrt{-\frac{p}{3}}$ and the root $\frac{3q}{p}$, where $p < 0$, (then $q^2 = -\frac{4p^3}{27} > 0$).

For $p > 0$ the relation $\frac{q^2}{4} + \frac{p^3}{27} > 0$ holds for an arbitrary q. If $p = 0$ and $q \neq 0$, the equation is of the form $x^3 + q = 0$, thus $x = \sqrt[3]{-q}$. The given equation has the unique real root. For $p = q = 0$ the equations is of the form $x^3 = 0$, it posesses one triple-root $x = 0$.

These results can be summarised in the following well-arranged table ([5], 511[6]):

	$x^3 + px + q = 0$
$\frac{q^2}{4} + \frac{p^3}{27} > 0$	1 real and 2 complex conjugated roots
$\frac{q^2}{4} + \frac{p^3}{27} < 0$	3 real different roots
$\frac{q^2}{4} + \frac{p^3}{27} = 0$	3 real roots, 1 of them is with the multiplicity greater than 1

Table 1.

This table could be more detailed and include also premises. The text above could be more concise.

At this moment the fact that the Cardan's expressions are of little value for calculating the roots is not necessarily so interesting. Perhaps we could reason using the discriminant $D = \left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3 \lessgtr 0$. But for $D > 0$ we have got roots in the imaginary form. The point here is not the fact that Cardan's expressions are unsuitable for this case. What is essential is that the impossibility to express real roots by real cube roots when $D > 0$ can be proved (see casus irreducubilis for cubic equations).

In the well-known mathematical tables ([1], 114[1]-115[3]), first of all, the general equation of order 3

$$Ax^3 + Bx^2 + Cx + D = 0, \quad A \neq 0,$$

is transformed to the normal form

$$x^3 + ax^2 + bx + c = 0,$$

then by substitution of $x = y - \frac{a}{3}$ to the reduced form. Now — see quotation:

Cardan' rule for reduced equation

$$\begin{aligned} y_1 &= u + v, \\ y_2 &= -\frac{u+v}{2} + \frac{u-v}{2} i\sqrt{3}, \\ y_3 &= -\frac{u+v}{2} - \frac{u-v}{2} i\sqrt{3}, \end{aligned}$$

where

$$u = \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}},$$

$$v = \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}.$$

Discriminant $D = \left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3$:
$D > 0$ determines one real and two complex conjugated roots.
$D = 0$ determines three real roots, one of them is the double root.

Casus irreducibilis

Discriminant $D < 0$ determines three real roots, which could be calculated using trigonometric functions (casus irreducibilis)

$$\begin{aligned} y_1 &= 2\sqrt{\frac{|p|}{3}}\cos\frac{\varphi}{3}, \\ y_2 &= -2\sqrt{\frac{|p|}{3}}\cos\left(\frac{\varphi}{3} - 60^\circ\right), \\ y_3 &= -2\sqrt{\frac{|p|}{3}}\cos\left(\frac{\varphi}{3} + 60^\circ\right); \end{aligned}$$

φ could be counted from the equation $\cos\varphi = \dfrac{-\frac{q}{2}}{\sqrt{\left(\frac{|p|}{3}\right)^3}}$.

The values of x found will be obtained from the substitution above $x = y - \frac{a}{3}$.

This well-arranged specification, typical for mathematical tables, is suitable for an immediate application, but it does not disclose the process of building and derivation of the given rules.

However, when writing a mathematical (nice to read) text we shouold be aware of who is it addressed to. Scientists — specialists have quite different needs than students studying mathematics; students of mathematics have quite different needs than students using mathematics occasionally only etc. A difference is also between students of lower and upper level of study. But everybody's requirement is the correct matter-of-fact, tabular, well-arranged and intelligible textbook. The quotation above shows that authors of older books remembered this fact but they did not make use of it often enough. However, the present time affords lots of technical instruments for them; and so we seek means to influence authors of (non) mathematical texts to use those instruments.

References

[1] H. J. Bartsch, Matematické vzorce, Praha, SNTL/ALFA 1971.

[2] D. Bittnerová, J. Vild, *Distance Principles for Full Time Studies*, in: VIth Czech-Polish Mathematical School, Litoměřice červen 1999, UJEP Ústí n. Labem 1999, 191-196.

[3] J. Vild, D. Bittnerová, *Destrukce lineární struktury textů*, in: Seminář o využití výpočetní techniky na technických VŠ, Liberec 1999, 23-28.

[4] L. Meertens, *Algorithmics — Towards programming as a mathematical activity*, in: Mathematics and Computer Science, Proceedings of the CWI symposium, Amsterdam 1983, 299-334.

[5] V. I. Smirnov, *Učebnice vyšší matematiky I*, Praha, Nakladatelství ČAV, 1954.

Technical University in Liberec
Faculty of Education
Hálkova 6
461 17 Liberec
Czech Republic
E-mail: daniela.bittnerova@vslib.cz