Mathematica Balkanica

Mathematical Society of South-Eastern Europe
A quarterly published by
the Bulgarian Academy of Sciences – National Committee for Mathematics

Mathematica Balkanica

New Series Vol. 9, 1995, Fasc. 2-3

New Analytic Solution of the Trinomial Algebraic Equation $z^n+pz+q=0$ by Means of the Goursat Hypergeometric Function, II

Pavel G. Todorov

Presented by V. Kiryakova

This paper is a continuation of the paper [1]. Now, for the cubic equation (1) reduced to the one-parametric form (3), a simple resolvent differential equation, representations of the analytic solution by elementary functions, new geometry of the roots, and a possibility for tabulation of the roots in the "irreducible case" are found.

1. Introduction

For the cubic equation

$$(1)\qquad z^3+pz+q=0 \qquad (p\neq 0),$$

our transformations [1,item 3.1]

$$(2)\begin{cases} z = \zeta\ {}_0\sqrt{-p/3} & \left(\arg {}_0\sqrt{-p/3} = \frac{\theta}{2};\ \arg(-p) = \theta;\ -\pi<\theta\leq\pi\right), \\ t = -\frac{q}{2({}_0\sqrt{-p/3})^3} & \left(\arg({}_0\sqrt{-p/3})^3 = 3\theta/2\right), \end{cases}$$

where the sign 0 denotes the principal value of the quadratic radical lying in the angle $(-\pi/2,\pi/2]$, reduced it to the following one-parametric form

$$(3)\qquad \zeta^3-3\zeta-2t=0 \qquad (\forall t).$$

The three-valued algebraic function $\zeta = f(t)$ of the equation (3) is inverse to the polynomial

$$t = \phi(\zeta) = \frac{1}{2}\zeta^3 - \frac{3}{2}\zeta. \tag{4}$$

According to (4) (see details in [1, item 3.1]), the three branches $\zeta_k = f_k(t)$ $(k = 0,1,2)$ of the function $\zeta = f(t)$ are regular and univalent in the fundamental region $G : \{t \notin g\}$, where $g = g_0 \cup g_1$ with $g_0 : [-\infty, -1]$ and $g_1 : [1, +\infty]$, and they are determinated by the values $\zeta^0_{0,1} = f_{0,1}(0) = \pm\sqrt{3}$ and $\zeta^0_2 = f_2(0) = 0$ (the branch $\zeta_0 = f_0(t)$ is the principal solution of the equation (3)).

2. Resolvent differential equation of the cubic equation

In [2, pp. 36–37 and 41] it is proved that the analytic solution of each trinomial algebraic equation of degree n in the Mellin form satisfies some superior Goursat hypergeometric differential equation of order n which is called Mellin resolvent differential equation. According to this result, Mellin resolvent differential equation of the cubic equation in the Mellin form is some Clausen hypergeometric differential equation of the third order (as the Goursat equations for $n = 3$ are called). Now, we shall show that the analytic solution of the cubic equation in our form (3) satisfies one special Gauss hypergeometric differential equation of the second order which we shall, in turn, call resolvent on the basis of the following propositions:

Theorem 1. *In the fundamental region G every two of the three branches $\zeta_k = f_k(t)$ $(k = 0,1,2)$ and the corresponding continuous continuations $\zeta_k^{\pm}$ $(k = 0,1,2)$ on the banks of the cut $g\backslash\{\pm 1\}$ of the three-valued algebraic function $\zeta = f(t)$ of the cubic equation (3) constitute a fundamental system of solutions of the special hypergeometric differential equation*

$$(1-t^2)\frac{d^2\zeta}{dt^2} - t\frac{d\zeta}{dt} + \frac{1}{9}\zeta = 0 \tag{5}$$

with three regular singular points $t = \pm 1$ and $t = \infty$.

Proof. The assertion follows from the results in [1], namely: both from Theorems 3 and 4 and from Theorem 6 (in a such easier way), in view of the same structure of the formulas (68)–(69) and the linearity of the argument in them. In the cut plane $z \notin [-\infty, 0] \cup [1, +\infty]$ the Gauss hypergeometric functions of the first and the second kind

$$(6)\qquad \begin{cases} u_2(1-2z) = F(-1/3, 1/3; 1/2; z), \\ u_3(1-2z) = 2/3\ {}_0\sqrt{z}\, F(1/6, 5/6; 3/2; z) \end{cases}$$

and their continuous continuation on the banks of the cut $z \in (-\infty, 0)$ are fundamental solutions of the hypergeometric equation of the same name

$$(7)\qquad z(1-z)\frac{d^2w}{dz^2} + (\frac{1}{2} - z)\frac{dw}{dz} + \frac{1}{9}w = 0$$

with three regular singular points $z = 0, 1$ and $z = \infty$.

By changing the variables $z = (1-t)/2$, $w(z) = \zeta(t)$ the equation (7) is reduced to the equation (5), while the functions (6) to the functions $u_{2,3}(t)$, considered in G. Hence, in the fundamental region G the functions $u_{2,3}(t)$ and their continuous continuations on the banks of the cut $g_1\backslash\{1\}$ are fundamental solutions of the equation (5).

Now it is obvious that in [1, Theorem 6] every two of the representations (68) and of the representations (69) (for $t \in g_1\backslash\{1\}$) are linear substitutions of the corresponding solutions $u_{2,3}(t)$ with determinants different from zero. Hence, conversely, in [1, Theorem 6] every two of the functions $\zeta_k = f_k(t)$ ($k = 0, 1, 2$) and of the functions $\zeta_k^{\pm}$ ($k = 0, 1, 2$) form a fundamental system of solutions of the equation(5) in the fundamental region G and on the infinite interval $g\backslash\{\pm 1\}$, respectively.

This completes the proof of Theorem 1.

Every two of the representations (48) from Theorem 3 in [1] are also linear substitutions of the other pair of Gauss hypergeometric functions of the first and the second kind $u_{0,1}(t)$ with determinants different from zero. Thence and from Theorem 1 it follows that, conversely, the pair $u_{0,1}(t)$ ($t \in G$) is also a fundamental system of solutions of the equation (5) (this can be shown by the substitution $z = t^2$ as well, which reduced (5) to the corresponding normal Gauss hypergeometric equation, differing from the equation (7) with the factor $1/36$ in front of w). Analogously, from the representations (76)–(77) in Theorem 7 in [1] it follows that each pair of functions ${}_k\sqrt[3]{2t}\, u_4(t)$, $(1/\,{}_k\sqrt[3]{2t})u_5(t)$ ($k = 0, 1, 2;\ 0 < |\arg t| < \pi$) is a fundamental system of solutions of (5). Besides, from the permanent principle of the solutions under their analytic continuations it follows that each of the branches $\zeta_0 = f_0(t)$ and $\zeta_1 = f_1(t)$, represented by the corresponding function (63) from Theorem 5 in [1], is a solution of the differential equation (5) everywhere in its own domain of existence.

According to Theorem 1, the roots of the cubic equation (3) as functions of the parameter t are expressed with linear substitutions of each pair of fundamental solutions of the differential equation(5), and, conversely, the solutions of the equation (5) are expressed with linear substitutions of any pair of roots of (3). Hence, the problems for the solution and the investigation of the algebraic equation (3) and of the differential equation (5) are equivalent. Because of that we shall call the equation (5) the resolvent differential equation of the cubic equation (3). Since the resolvent equation (5) has an infinite set of fundamental systems of solutions, it follows that the roots of the cubic equation (3) have an infinitre set of such representations as well. In particular, we can express the roots by other special functions. For that purpose, in the resolvent equation (5) a corresponding transformation should be made. We shall restrict ourselves to obtain representations of the roots in elementary functions.

3. Representation of the branches by elementary functions

The following theorem is a bridge between the analytic solution of the cubic equation (3) and the algebraic solution by the Cardano classical formula:

Theorem 2. *Let the fundamental region G be mapped by the principal and regular branch of the function*

$$(8)\qquad \tau(t) = \arccos t = \frac{1}{i}\ln(t+{}_0\sqrt{t^2-1})$$

univalently onto the band $\{0<\Re\tau<\pi\}$, where the upper bank of the cut g_1 and the lower bank of the cut g_0 are respectively mapped by means of the functions

$$(9)\qquad \begin{cases} \tau(t+i0) = \arccos(t+i0) = \frac{1}{i}\ln(t+{}_+\sqrt{t^2-1}) & (t\in g_1),\\ \tau(t-i0) = \arccos(t-i0) = \frac{1}{i}\ln(t+{}_+\sqrt{t^2-1}) & (t\in g_0) \end{cases}$$

one-to-one on the right bank of the cut $\{\Re\tau=0;\ \Im\tau\le 0\}$ and on the left bank of the cut $\{\Re\tau=\pi;\ \Im\tau\ge 0$.

Then the regular and univalent branches and the boundary values of the three-valued algebraic function $\zeta=f(t)$ of the cubic equation (3) have representations in elementary trigonometric functions

$$(10)\qquad \begin{cases} \zeta_k = f_k(t) = 2\cos\frac{\arccos t+2k\pi}{3} & (k=0,1,2;\ t\in G),\\ \zeta_k^+ = 2\cos\frac{\arccos(t+i0)+2k\pi}{3} & (k=0,1,2;\ t\in g_1),\\ \zeta_k^- = 2\cos\frac{\arccos(t-i0)+2k\pi}{3} & (k=0,1,2;\ t\in g_0). \end{cases}$$

Corollary. *Let the one-valued function*

$$\tau \;=\; \arccos t \;=\; \arg(t + {}_0\sqrt{t^2-1}) \;-\; i\,\ln|t + {}_0\sqrt{t^2-1}| \tag{11}$$

be introduced for consideration in the finite t-plane, the set of values of which constitute the band $\{0 < \Re\tau < \pi\}$ and the rays $\{\Re\tau = 0;\ \Im\tau \le 0\}$,$\{\Re\tau = \pi;\ \Im\tau \ge 0\}$, joined to it.

Then the roots z_k $(k = 0, 1, 2)$ of the cubic equation (1) are equal to

$$z_k \;=\; 2\,{}_0\sqrt{-p/3}\ \cos\frac{\arccos t + 2k\pi}{3} \qquad (k = 0, 1, 2;\ t \;=\; -\frac{q}{2({}_0\sqrt{-p/3})^3};\ p \neq 0). \tag{12}$$

P r o o f. By the transformation $t = \cos\tau$, $\zeta(\tau) = \eta(\tau)$ the resolvent equation (5) takes the form

$$\frac{d^2\eta}{d\tau^2} + \frac{1}{9}\eta \;=\; 0, \tag{13}$$

which has the functions $\cos\tau/3$ and $\sin\tau/3$ as fundamental solutions. Hence, the resolvent equation (5) tolerates, as fundamental solutions, the elementary functions

$$u_6(t) \;=\; \cos\frac{\mathrm{Arccos}t}{3}, \qquad u_7(t) \;=\; \sin\frac{\mathrm{Arccos}t}{3} \tag{14}$$

for an arbitrary regular branch of the many-valued function Arccos$t$.

Let us choose the regular branch (8), constituted by the regular branch of the Žukovski inverse function

$$v(t) \;=\; t + {}_0\sqrt{t^2-1} \quad (t \in G), \tag{15}$$

the boundary values of which are given by the relations

$$\begin{aligned} v(t \pm i0) &= t \mp {}_+\sqrt{t^2-1}\ \ (t \in g_0),\\ v(t \mp i0) &= t \mp {}_+\sqrt{t^2-1}\ \ (t \in g_1), \end{aligned} \tag{16}$$

obtained by circuit of the branch points $t = \mp 1$ along the indicated banks of the cut g. Besides, the upper and the lower bank of the cut g_0 and the lower and the upper bank of the cut g_1 are mapped by means of the corresponding functions (16) one-to one onto the upper banks of the segments $v \in [-\infty, -1]$, $v \in [-1, 0]$, $v \in [0, 1]$ and $v \in [1, +\infty]$, respectively. Hence, the function (15) maps G univalently onto the upper plane $\Im v > 0$ in the following way: the half-plane $\Im t > 0$ onto the half-disc $|v| > 1 \cap \Im v > 0$, the segment $t \in (-1, 1)$ onto the half-circle $|v| = 1 \cap \Im v > 0$ and the half-plane $\Im t < 0$ onto the half-disc $|v| < 1 \cap \Im v > 0$. Let us note that the function (15) maps the imaginary axis $t \in (-i\infty, +i\infty)$ onto the imaginary half-axis $v \in (0, +i\infty)$. By means of the functions (16) and (15), the principal branch of the function (8) maps the two banks of the cuts g_0 and g_1 one-to-one onto the left and the right bank of the cuts $\Re\tau = \pi$ and $\Re\tau = 0$, respectively, and the region G univalently onto the band $\{0 < \Re\tau < \pi\}$.

Considering the functions (14) chosen so as, to express the other pair of fundametal solutions (49) (for $t \in G$) in [1, Theorem 3] of the resolvent equation (5) by their linear substitution

$$(17) \qquad u_j(t) = c_{6j}u_6(t) + c_{7j}u_7(t) \qquad (j = 0, 1),$$

where c_{6j} and c_{7j} are some constants. Differentiating (17) with the help of the rule for differentiating the hypergeometric function, we obtain two systems for their determination, which for $t = 0$ yield $c_{60} = -c_{71} = \sqrt{3}/2$ and $c_{70} = c_{61} = 1/2$. Thus, from (17) we find that the Gauss hypergeometric functions of the first and the second kind have the following representations in elementary functions:

$$(18) \qquad \begin{cases} u_0(t) = \sin \frac{\pi + \arccos t}{3}, \\ u_1(t) = \cos \frac{\pi + \arccos t}{3}, \end{cases} \qquad (t \in G),$$

where the property of $u_o(t)$ of being even and the property of $u_1(t)$ of being odd become evident by means of the equality $\arccos t + arccos(-t) = \pi$, valid on the banks of the cut g as well. Putting (18) in (48) in [1, Theorem 3], we obtain the first formula from (10) for the functions $\zeta_k = f_k(t)$ $(k = 0, 1, 2)$. The continuous continuation of these functions on the banks of the cut g is realized by the formulas (54)–(55) in [1] and the superpositions $\arccos(t \pm i0)$ $(t \in g)$ of the logarithmic function and the corresponding functions (16). Choosing, for example, the functions $v(t + i0)$ for $t \in g_1$ and $v(t - i0)$ for $t \in g_0$, i.e. the functions (9), we obtain other two formulas from (10) for the boundary values $\zeta_k^{\pm}$ $(k = 0, 1, 2)$. For transition of one bank of the cut g to another, we have here the functional relations

$$(19) \qquad \begin{cases} \arccos(t - i0) = -\arccos(t + i0) & (t \in g_1), \\ \arccos(t + i0) = 2\pi - \arccos(t - i0) & (t \in g_0), \end{cases}$$

by means of which, conversely, we can immediately obtain the functional relations (54)–(55) in [1].

This completes the proof of Theorem 2 and its Corollary.

Comparing (68) in [1, Theorem 6] with the first formula from (10), we obtain the following representations in elementary functions for the other pair of Gauss hypergeometric functions of the first and the second kind

$$\text{(20)} \qquad \begin{cases} u_2(t) = \cos \frac{\arccos t}{3}, \\ u_3(t) = \sin \frac{\arccos t}{3} \end{cases} \qquad (t \in G)$$

as well. The continuous continuation of the trigonometric functions from (18) and (20) on the banks of the cut g is obtained immediately by the formulas (9) and (19). In particular, continuing continuously the second formula from (20) on the upper bank of the cut g_1 and applying (74) in [1], we obtain

$$\text{(21)} \qquad u_3(t) = -\sin \frac{\arccos(t + i0)}{3} \qquad (t \in g_1).$$

Entering formula (21) in formula (75) in [1], we obtain the jumps of the branches by the transition through the cut g in trigonometric functions.

With the help of the formula for cos of a triple angle, we can make a verification of the roots (12) of the cubic equation (1). Conversely, by the formula for cos of a triple angle we can obtain the formula (12) from the equations (1)–(3) themselves, but without explicitly and exactly introduced function (11) it would have had formal character only. In particular, when the cubic equation (1) is with real coefficients and is algebraicly irreducible, i. e. for $t \in (-1, 1)$ and $p < 0$ (see [1, item 3.2] and Theorem 4 below), then the formula (12) is known, and in the literature it is given as a substitute of the Cardano formula for calculating the roots. From our theory it follows that the formula (12) is already valid for an arbitrary complex t, i. e. for an arbitrary cubic equation (1).

Theorem 2 gives the possibility of finding explicit algebraic representations of the branches and of the boundary values of the three-valued algebraic function $\zeta = f(t)$. This, in its turn, generates a convenient representation for application of the roots (12) by real elementary functions of the arbitrary complex parameter t.

Theorem 3. *Let the fundamental region G be mapped univalently by the three regular branches of the cubic radical*

$$\text{(22)} \quad a_k(t) = {}_k\sqrt[3]{t + {}_0\sqrt{t^2 - 1}} = a(t) \exp i \frac{\alpha(t) + 2k\pi}{3} \qquad (k = 0, 1, 2),$$

respectively, onto the angles $2k\pi/3 < \arg a_k(t) < (2k+1)\pi/3$ $(k = 0,1,2)$, *where the upper bank of the cut* g_1 *and the lower bank of the cut* g_0 *are mapped one-to-one by means of the functions*

$$(23)\quad \begin{cases} a_k(t+i0) = {}_+\sqrt[3]{t + {}_+\sqrt{t^2-1}} = a(t)\exp i\frac{2k\pi}{3} & (k=0,1,2; t\in g_1), \\ a_k(t-i0) = {}_+\sqrt[3]{t + {}_+\sqrt{t^2-1}} = a(t)\exp i\frac{(2k+1)\pi}{3} & (k=0,1,2; t\in g_0), \end{cases}$$

respectively, onto the interior banks of the cuts $[\exp i2k\pi/3, \infty\cdot\exp i2k\pi/3]$ *and* $[0, \exp i(2k+1)\pi/3]$ $(k = 0,1,2)$ *along the sides of these angles, where by* $a(t)$ *and* $\alpha(t)$ *the real functions*

$$(24)\begin{cases} a(t) = {}_+\sqrt[3]{|t + {}_0\sqrt{t^2-1}|}, \\ \alpha(t) = \arg(t + {}_0\sqrt{t^2-1}) = \arccos\frac{1}{2}(|t+1| - |t-1|) \quad (0 \overset{t\in g_1}{\le} \alpha(t) \overset{t\in g_0}{\le} \pi) \end{cases}$$

of the arbitrary complex argument t *are denoted.*

Then, the regular and univalent branches and the boundary values of the three-valued algebraic function $\zeta = f(t)$ *of thye cubic equation (3) have representations by elementary algebraic and real trigonometric functions*

$$(25)\quad \begin{cases} \zeta_k = f_k(t) = a_k(t) + \frac{1}{a_k(t)} = \\ = r(t)\cos\frac{\alpha(t)+2k\pi}{3} + i\rho(t)\sin\frac{\alpha(t)+2k\pi}{3} \end{cases} \quad (k = 0,1,2; t\in G),$$

$$(26)\quad \begin{cases} \zeta_k^+ = a_k(t+i0) + \frac{1}{a_k(t+i0)} = \\ = r(t)\cos\frac{2k\pi}{3} + i\rho(t)\sin\frac{2k\pi}{3} \end{cases} \quad (k=0,1,2; t\in g_1),$$

$$(27)\quad \begin{cases} \zeta_k^- = a_k(t-i0) + \frac{1}{a_k(t-i0)} = \\ = r(t)\cos\frac{(2k+1)\pi}{3} + i\rho(t)\sin\frac{(2k+1)\pi}{3} \end{cases} \quad (k = 0,1,2; t\in g_0),$$

where $r(t)$ *and* $\rho(t)$ *denote the real functions of the complex argument* t:

$$(28)\qquad r(t) = a(t) + \frac{1}{a(t)}, \quad \rho(t) = a(t) - \frac{1}{a(t)} \quad (\forall t).$$

Corollary 1. *Let for all finite t the set, constituted by the three values of the cubic radical (22) by means of (24), be introduced for consideration.*

Then the roots (12) of the cubic equation (1) are also equal to

$$(29)\quad \begin{cases} z_k = {}_0\sqrt{-p/3}\left(a_k(t) + \frac{1}{a_k(t)}\right) = & (k=0,1,2;\ t=-\frac{q}{2({}_0\sqrt{-p/3})^3};\ p\neq 0) \\ \quad = {}_0\sqrt{-p/3}\left(r(t)\cos\frac{\alpha(t)+2k\pi}{3} + i\rho(t)\sin\frac{\alpha(t)+2k\pi}{3}\right). \end{cases}$$

Corollary 2 (The Cardano formula). *Let for all finite t, besides the set of the condition of Corollary 1, the set of the reciprocal values*

$$(30)\quad b_k(t) = \frac{1}{a_k(t)} = {}_k\sqrt[3]{t - {}_0\sqrt{t^2-1}} = b(t)\exp i\frac{\beta(t)-2k\pi}{3} \quad (k=0,1,2)$$

be introduced for consideration as well, where

$$(31)\quad \begin{cases} b(t) = \frac{1}{a(t)} = +\sqrt[3]{|t - {}_0\sqrt{t^2-1}|}, \\ \beta(t) = -\alpha(t) = \arg(t - {}_0\sqrt{t^2-1}) \quad (-\pi \overset{t\leq -1}{\leq} \beta(t) \overset{t\geq 1}{\leq} 0). \end{cases}$$

Then the roots (29) of the cubic equation (1) are represented by means of these two sets by the Cardano formula in the form

$$(32)\quad \begin{cases} z_k = {}_0\sqrt{-p/3}(a_k(t)+b_k(t)) = & (k=0,1,2;\ t=-\frac{q}{2({}_0\sqrt{-p/3})^3};\ p\neq 0) \\ \quad = {}_0\sqrt{-p/3}\left[(a_0(t)+b_0(t))\cos\frac{2k\pi}{3} + i\,(a_0(t)-b_0(t))\sin\frac{2k\pi}{3}\right]. \end{cases}$$

Proof. The formula (25) follows from the first formula in (10) if we express all cos by the Euler formula and apply the formula (8) for the corresponding arccos (for $k = 0$ the principal branch (25) is constituted by the principal branch of the radical (22)). Hence, each pair of radicals $a_k(t)$, $1/a_k(t)$ ($k = 0,1,2$) is a fundamental system of solutions of the resolvent differential equation (5). The continuous continuation of the formulas (25) on the banks of the cut g is realized by the formulas (54)–(55) in [1] for the left-hand side, and for the right-hand side by the superpositions $a_k(t \pm i0)$ ($k = 0,1,2;\ t \in g$) of the cubic radical (22) and the corresponding functions (16). Choosing, for example, $a_k(t+i0)$ for $t \in g_1$ and $a_k(t-i0)$ for $t \in g_0$, i.e. the functions (23), we obtain the formulas (26)–(27). For transition of one bank of the cut g to another, the functional relations

$$(33)\qquad a_k(t+i0)a_k(t-i0) = \begin{cases} \exp i\frac{4k\pi}{3}, & t\in g_1, \\ \exp i\frac{2(2k+1)\pi}{3}, & t\in g_0, \end{cases} \qquad (k=0,1,2)$$

are valid here, by means of which, conversely, we can obtain the functional relations (54)–(55) in [1].

According to the noted characteristics of the mapping of the region G by the function (15) and of the upper (lower) bank of the cut g_1 (g_0) by the corresponding functions (16), for the modulus $a(t)$ from (24) we have the following characteristic: $a(t)>1$ for $0\le \arg(t-1)<\pi$, $a(t)=1$ for $t\in[-1,1]$ and $0<a(t)<1$ for $-\pi\le\arg(t+1)<0$, i. e. for the functions $r(t)$ and $\rho(t)$ from (28) we have: $r(t)\ge 2$ and $-\infty<\rho(t)<+\infty$, where $r(t)>|\rho(t)|$. Besides, $a(t)\to+\infty$ ($r(t)\to+\infty$, $\rho(t)\to+\infty$) and $a(t)\to 0$ ($r(t)\to+\infty$, $\rho(t)\to-\infty$) for $t\to\infty$ in an arbitrary direction in the upper half-plane $0\le\arg(t-1)<\pi$ and in the lower half-plane $-\pi\le\arg(t+1)<0$, respectively.

Now, we shall find an explicit form of the argument $\alpha(t)$ of the Žukovski inverse function (15). For that purpose we shall use the identities

$$(34)\qquad \begin{cases} |t+1|+|t-1| = |t+{}_0\sqrt{t^2-1}|+|t-{}_0\sqrt{t^2-1}|, \\ \Re t = \frac{1}{4}\left(|t+1|^2-|t-1|^2\right), \end{cases} \qquad (\forall t),$$

and the ensuing from the Žukovski function (15) itself identity

$$(35)\qquad t=\frac{1}{2}\left(v(t)+\frac{1}{v(t)}\right)\qquad (v(t)=|v(t)|e^{i\alpha(t)};\ t\in G).$$

From (35) it follows that

$$(36)\qquad \Re t=\frac{1}{2}\left(|v(t)|+\frac{1}{|v(t)|}\right)\cos\alpha(t),$$

whence by means of (34) we find the formula

$$(37)\qquad \cos\alpha(t)=\frac{1}{2}\left(|t+1|-|t-1|\right)\qquad (t\in G),$$

valid, according to (16), on the banks of the cut g as well. By the triangle inequality it can be verified that the values of the right-hand side of (37) really do not go out of the segment $[-1,1]$. Taking into consideration that $0 \le \alpha(t) \le \pi$, from (37) we obtain the formula for $\alpha(t)$ in (24).

This completes the proof of Theorem 3 and its Corollaries.

With the help of Theorem 3, from (54)–(55) in [1] we obtain the jumps of the branches by the transition through the cut g in radicals:

$$(38)\qquad \begin{cases} f_{1,2}(t+i0) - f_{1,2}(t-i0) = \pm i\sqrt{3}\rho(t) & (t \in g_1), \\ f_{0,2}(t+i0) - f_{0,2}(t-i0) = \mp i\sqrt{3}\rho(t) & (t \in g_0). \end{cases}$$

The connection between these relations is realized by the equality $\rho(-t) = -\rho(t)$, ensuing from (28). The comparison of (38) with [1, formulas (62) and (75)] yields the corresponding equalities.

Comparing the formulas (48) and (68) from Theorems 3 and 6 in [1] with formula (25) from Theorem 3, we obtain the following representations in radicals of the considered pair of Gauss hypergeometric functions of the first and the second kind:

$$(39)\qquad \begin{cases} u_0(t) = -\frac{1}{2i}\left(a_2(t) - \frac{1}{a_2(t)}\right), \\ u_1(t) = -\frac{1}{2}\left(a_2(t) + \frac{1}{a_2(t)}\right), \end{cases} \qquad (t \in G),$$

$$(40)\qquad \begin{cases} u_2(\pm t) = \pm\frac{1}{2}\left(a_{0,1}(t) + \frac{1}{a_{0,1}(t)}\right), \\ u_3(\pm t) = \frac{1}{2i}\left(a_{0,1}(t) - \frac{1}{a_{0,1}(t)}\right), \end{cases} \qquad (t \in G),$$

where the upper (lower) signs correspond to the first (second) indices. The continuous continuation of these formulas on the banks of the cut g is obtained immediately by the formulas (23) and (33).

Analogously, from the system obtained by the comparison of the formulas (76)–(77) from Theorem 7 in [1] with the formula (25) from Theorem 3, representations of the functions (79) in [1, Theorem 7] in the upper and lower t-plane in radicals

$$(41)\qquad \begin{cases} {}_k\sqrt[3]{2t}\,u_4(t) = a_k(t), \ \frac{1}{{}_k\sqrt[3]{2t}}u_5(t) = \frac{1}{a_k(t)} & (k=0,1,2; 0<\arg t<\pi), \\ \epsilon^k\,{}_k\sqrt[3]{2t}\,u_4(t) = \frac{1}{a_k(t)}, \ \frac{1}{\epsilon^k\,{}_k\sqrt[3]{2t}}u_5(t) = a_k(t) & (k=0,1,2; -\pi<\arg t<0), \end{cases}$$

follow. From (41) the identity (82) in [1] for $0<|\arg t|<\pi$ immediately follows, whence, by means of the analytic continuation we establish that it is true in the domain $t \notin [-1,1]$ as well, to which we can also add the boundary points $t = \pm 1$.

Remark 1. From the formulas for cos and sin of a triple angle, the identities

$$(42) \qquad \begin{cases} \cos\alpha(t) &= 4\prod_{k=0}^{2}\cos\frac{\alpha(t)+2k\pi}{3}, \\ \sin\alpha(t) &= -4\prod_{k=0}^{2}\sin\frac{\alpha(t)+2k\pi}{3} \end{cases} \qquad (\forall t)$$

follow. With their help we establish that the argument $\alpha(t)$ (its boundary values included) of the Žukovski inverse function (15) is connencted with the real and imaginery parts of the roots $\zeta_k = \xi_k(t)+i\eta_k(t)$ $(k = 0,1,2)$ of the cubic equation (3), taken from the factor in the parentheses of the second formula in (29), by the formulas

$$(43) \qquad \begin{cases} \cos\alpha(t) &= \frac{4}{r^3(t)}\,\xi_0(t)\xi_1(t)\xi_2(t) \qquad (\forall t), \\ \sin\alpha(t) &= -\frac{4}{\rho^3(t)}\,\eta_0(t)\eta_1(t)\eta_2(t) \quad (t \notin [-1,1]). \end{cases}$$

Remark 2. The second formula (37) obtained for the argument $\alpha(t)$ of the Žukovski inverse function (15) gives us a possibility of writing an explicit form for the inverse trigonometric function (8) (respectively, (11)). More generally, for the two basic regular and univalent branches $(\arccos t)^{\pm}$ of the classical many-valued function $\operatorname{Arccos} t$ we obtain the explicit formula

$$(44) \qquad \begin{cases} (\arccos t)^{\pm} = \frac{1}{i}\ln(t \pm {}_0\sqrt{t^2-1}) = \\ = \pm\arccos\frac{1}{2}(|t+1|-|t-1|) - i\ln|t \pm {}_0\sqrt{t^2-1}| \end{cases} \qquad (t \in G),$$

where the values of the real arccos on the right-hand side lie on the segment $[0,\pi]$. Formula (44) is valid on the lower bank of the cut g_0 and the upper bank of the cut g_1 as well (it is also valid on the opposite banks of these cuts if the signs $\pm$ in front of the radicals are replaced with the signs $\mp$). An arbitrary regular and univalent branch of the function $(\operatorname{Arccos} t)^{\pm}$ is obtained if we add $2\nu\pi$ $(\nu = 0,\pm1,\pm2,\dots)$ to the corresponding basic branch (44). The values of $(\operatorname{Arccos} t)^{\pm}$ lie on the bands $2\nu\pi < \Re(\operatorname{Arccos} t)^{+} < (2\nu+1)\pi$ and $(2\nu-1)\pi < \Re(\operatorname{Arccos} t)^{-} < 2\nu\pi$ $(\nu = 0,\pm1,\pm2,\dots)$, respectively.

Remark 3. The formulas (25)–(27) show that they can be obtained in an artificial way from the cubic equation (3) setting $\zeta = \delta e^{i\varphi} + 1/\delta e^{i\varphi}$, where the modulus δ and the argument φ have to be determined. Thus, we obtain the quadratic equation $u + (1/u) - 2t = 0$, where $u = \delta^3 e^{3i\varphi}$. Comparing this expression, in which we choose $\delta = a(t)$, with the corresponding root, we obtain $\varphi_k = (\alpha(t)+2k\pi)/3$ $(k = 0,1,2)$.

Remark 4. The inverse route from the Cardano formula (32) to Theorem 3 in [1] is also possible. In order to do this, $t = \cos\tau$ should be set in (32) and, further, each stage should be passed in the already explained manner. By the inverse route we shall obtain Theorem 3 in [1] independently of the basic Theorem 1 in [1], being its corollary for $n = 3$. Of course, the inverse approach has a local character since it refers to the cubic equation, whose algebraic solution (the Cardano formula) is known to us.

As an application of the second formula from (29), we shall calculate the magnitudes $\alpha(t)$, $a(t)$, $r(t)$ and $\rho(t)$ by formulas (24) and (28) in the case when the coefficients $p \neq 0$ and q of the cubic equation (1) are real. For an arbitrary real q it is possible that $p<0$ or $p>0$, so the parameter t, calculated by (2) (or by (29)) is a real or a pure imaginary number. If $p<0$, for all $t \in [-1,1]$, i. e. for $-2({}_{+}\sqrt{-p/3})^3 \leq q \leq 2({}_{+}\sqrt{-p/3})^3$, we have $\alpha(t) = \arccos t$, $a(t) = 1$, $r(t) = 2$ and $\rho(t) = 0$, by which the second formula from (29) is reduced to the formula (12). If $p<0$, for all $t \geq 1$ $(q \leq -2({}_{+}\sqrt{-p/3})^3)$ or $t \leq 1$ $(q \geq 2({}_{+}\sqrt{-p/3})^3)$ we have $\alpha(t) = 0$ and $\alpha(t) = \pi$, respectively, while the modulus $a(t)$ is given immediately by (24). If $p > 0$, for all $t \in (-i\infty, +i\infty)$ $(-\infty < q < +\infty)$ we have $\alpha(t) = \pi/2$, while the modulus $a(t)$ is determined by (24). In general, the second formula from (29) as a new formula for the numerical solution of an arbitrary cubic equation (1) generalizes in real functions of the arbitrary complex parameter t the classical real formula (12) for the "irreducible case" $t \in (-1,1)$.

Besides, from the comparison of Theorem 3 and Corollary 1 with Corollary 2 it is obvious that our formulas (25)–(27) and (29) are more suitable for investigations than the Cardano formula (32). Thus, for example, the second formula from (29) yields a possibility of building completely the geometry of the roots of the cubic equation (1) which has not been made until now.

4. Geometry of the roots

The following theorem refers to the general position of the roots in the z-plane:

Theorem 4. *For an arbitrary finite t the three roots z_k $(k = 0,1,2)$ of the cubic equation (1) cannot lie on one straight line, except in the exceptional case for $t \in [-1,1]$ when these roots lie on the segment $z \in [-2\,{}_{0}\sqrt{-p/3}, 2\,{}_{0}\sqrt{-p/3}]$.*

Proof. It is sufficient to prove the theorem for the cubic equation (3). Taking its roots ζ_k $(k = 0,1,2)$ from the expressions in the parentheses of the second formula in (29), for their differences we obtain

$$(45)\begin{cases} \zeta_k - \zeta_{k+1} = & (k=0,1,2;\zeta_3=\zeta_0) \\ (-1)^k\sqrt{3}\left(r(t)\sin\frac{\alpha(t)-(k-1)\pi}{3} - i\rho(t)\cos\frac{\alpha(t)-(k-1)\pi}{3}\right) \end{cases}$$

Let us now constitute the affine ratio of the three roots and investigate its reality. For the imaginary part we obtain

$$(46)\quad \Im(\zeta_2,\zeta_0,\zeta_1) = \Im\frac{\zeta_1-\zeta_2}{\zeta_1-\zeta_0} = -\frac{\sqrt{3}}{2}\cdot\frac{r(t)\rho(t)}{r^2(t)\sin^2\frac{\alpha(t)+\pi}{3}+\rho^2(t)\cos^2\frac{\alpha(t)+\pi}{3}}\,.$$

For finite t the right-hand side can become zero only for $\rho(t) = 0$, i.e. for $a(t) = 1$, according to (28) and (24), which implies $|v(t)| = 1$, according to (15), where $\Im v(t) \overset{t=\mp 1}{\geq} 0$. But the original of the upper unit half-circle by the mapping (15) is the segment $t \in [-1,1]$, in which case the roots ζ_k $(k = 0,1,2)$ are real and lie on the segment $\zeta \in [-2,2]$.

This completes the proof of Theorem 4.

According to the classical Van der Berg theorem (see, for example, Marden's monograph [3], p. 9, Theorem (4.1) and pp. 12–13 for a detailed bibliography concerning the history of this theorem), the roots of the derivative equation of the cubic equation (1) lie at the foci of the ellipse, inscribed in the triangle, constructed by the three roots z_k $(k = 0,1,2)$ of the given equation, and touching the middles of its sides. The Van der Berg theorem does not give the conditions under which the triangle $z_0z_1z_2$ is not degenerate and the explicit equation of the ellipse. According to our Theorem 4, the Van der Berg theorem has non-degenerate meaning if $t \notin [-1,1]$ only. Now we shall find the equation of the ellipse as well and we shall also prove the Van der Berg theorem itself by a new and simple method with the help of the formula (29). In addition, the formula (29) yields the possibility of finding the equation of the circle, determined by the three roots z_k $(k = 0,1,2)$, and to establish the new fact that these roots lie on another ellipse with a special position. Geometrically this means that the roots of the cubic equation (1) for $t \notin [-1,1]$ are points of intersection of two completely determined curves of second degree. Let us formulate all these facts exactly:

Theorem 5. *For each finite $t \notin [-1,1]$ the non-degenerate triangle $z_0 z_1 z_2$, where z_0, z_1, z_2 are the roots of the cubic equation (1), is inscribed in the circle*

$$\left|z - \zeta^* {}_0\sqrt{-p/3}\right| = R {}_+\sqrt{|p|/3} \tag{47}$$

$$\left(\zeta^* = \frac{\cos\alpha(t)}{r(t)} + i\frac{\sin\alpha(t)}{\rho(t)};\ R = \frac{2 {}_+\sqrt{|t^2-1|}}{r(t)|\rho(t)|}\right)$$

and also in the ellipse

$$\left|z + 2 {}_0\sqrt{-p/3}\right| + \left|z - 2 {}_0\sqrt{-p/3}\right| = 2r(t) {}_+\sqrt{|p|/3} \tag{48}$$

with a large half-axis $r(t) {}_+\sqrt{|p|/3}$, a small half-axis $|\rho(t)| {}_+\sqrt{|p|/3}$ and foci $\mp 2 {}_0\sqrt{-p/3}$ and is circumscribed around the similar and identically-situated ellipse

$$\left|z + {}_0\sqrt{-p/3}\right| + \left|z - {}_0\sqrt{-p/3}\right| = r(t) {}_+\sqrt{|p|/3} \tag{49}$$

with a large half-axis $(1/2)r(t) {}_+\sqrt{|p|/3}$, a small half-axis $(1/2)|\rho(t)| {}_+\sqrt{|p|/3}$, foci at the points $\mp {}_0\sqrt{-p/3}$, which are the roots of the derivative equation of the given equation, touching the middles of its sides.

Proof. It is sufficient to prove the theorem for the roots ζ_k $(k = 0,1,2)$ of the cubic equation (3). According to Theorem 4 for $t \notin [-1,1]$, the roots ζ_k do not lie on a straight line and they determine a certain circle $|\zeta - \zeta^*| = R$. Its center ζ^* and radius R are determined by the system $(\zeta_k - \zeta^*)(\overline{\zeta}_k - \overline{\zeta}^*) = R^2$ $(k = 0,1,2)$, where ζ_k are taken from the second formula in (29). From here we can obtain a system of two equations with respect to ζ^* and $\overline{\zeta}^*$, whence, by means of (45), we come to the expression for ζ^* and by this for R in (47) as well. By means of (43) the formula obtained for ζ^* can be immediately expressed by the real and imaginary parts of the roots $\zeta_k = \xi_k + i\eta_k$ $(k = 0,1,2)$.

Further, by the second formula in (29) it follows immediately that the roots $\zeta_k = \xi_k + i\eta_k$ $(k = 0,1,2)$ lie on the ellipse from the ζ-plane

$$\frac{\xi^2}{r^2(t)} + \frac{\eta^2}{\rho^2(t)} = 1 \tag{50}$$

with half-axes $r(t)$ and $|\rho(t)|$ and foci ∓ 2, which in the z-plane has the complex equation (48), according to the first transformation from (2).

Now, we shall give a new and simple proof of the Van der Berg theorem for the non-degenerate case and we shall also find explicitly the ellipse (49). From the relation $\zeta_0 + \zeta_1 + \zeta_2 = 0$ it follows that the middles of the sides of the triangle $\zeta_0\zeta_1\zeta_2$ opposite the corresponding vertices are $-\zeta_k/2 = -\xi_k/2 - i\eta_k/2$ $(k = 0, 1, 2)$, and, analogously, in consequence of the second formula in (29), lie on the ellipse

$$\frac{\xi^2}{\frac{1}{4}r^2(t)} + \frac{\eta^2}{\frac{1}{4}\rho^2(t)} = 1 \tag{51}$$

with half-axes $(1/2)r(t)$ and $(1/2)|\rho(t)|$ and foci ∓ 1, which in the z-plane is given by (49). Now it can be verified that the pair of points $\zeta_{1,2}$, $\zeta_{2,0}$ and $\zeta_{0,1}$ lie on the tangents of the ellipse (51) at the points $-\zeta_k/2$ $(k = 0, 1, 2)$, respectively, i. e. the ellipse (51) is inscribed in the triangle $\zeta_0\zeta_1\zeta_2$. The points ∓ 1 at which its foci lie are the roots of the derivative equation of the equation (3).

This completes the proof of Theorem 5.

In Cartesian coordinates, the equation of the circle (47) can be written immediately. In particular, when the cubic equation (1) is with real coefficients $p \neq 0$ and q for which $t \notin [-1, 1]$ (see the corresponding cases in the end of item 3), the center of the circle (47) lies on the real axis, the origin excluded.

In Cartesian coordinates, the central equations of the homothetic ellipses (48) and (49) are $(p = p_1 + ip_2 \neq 0;\ t \notin [-1, 1];\ z = x + iy)$

$$\begin{aligned}&[\,|p|\,(r^2(t)-2) + 2p_1\,]x^2 + 4p_2xy + \\ &+ [\,|p|\,(r^2(t)-2) - 2p_1\,]y^2 - \frac{|p|^2}{3}r^2(t)\rho^2(t) = 0,\end{aligned} \tag{52}$$

$$\begin{aligned}&[\,|p|\,(r^2(t)-2) + 2p_1\,]x^2 + 4p_2xy + \\ &+ [\,|p|\,(r^2(t)-2) - 2p_1\,]y^2 - \frac{|p|^2}{12}r^2(t)\rho^2(t) = 0,\end{aligned} \tag{53}$$

with vertices $\pm r(t)\,{}_0\sqrt{-p/3}$, $\pm i\,|\rho(t)|\,{}_0\sqrt{-p/3}$ and $\pm(1/2)\,r(t)\,{}_0\sqrt{-p/3}$, $\pm(i/2)\,|\rho(t)|\,{}_0\sqrt{-p/3}$, respectively. In particular, for a real $p \neq 0$ $(t \notin [-1, 1])$ the ellipses (52) and (53) have central axial (canonical) equations

$$\frac{x^2}{\frac{|p|}{3}r^2(t)} + \frac{y^2}{\frac{|p|}{3}\rho^2(t)} = 1 \ (p<0), \qquad \frac{x^2}{\frac{p}{3}\rho^2(t)} + \frac{y^2}{\frac{p}{3}r^2(t)} = 1 \ (p>0), \tag{54}$$

and

$$\frac{x^2}{\frac{|p|}{12}r^2(t)} + \frac{y^2}{\frac{|p|}{12}\rho^2(t)} = 1 \ (p<0), \qquad \frac{x^2}{\frac{p}{12}\rho^2(t)} + \frac{y^2}{\frac{p}{12}r^2(t)} = 1 \ (p>0), \tag{55}$$

respectively.

Theorem 5 yields a possibility of making further investigations. That is why we shall again use the ζ-plane which is immediately transformed into z-plane by a rotation and a homothety with respect to the origin, according to the first formula from (2). Comparing the equations of the tangents of the circle $|\zeta - \zeta^*| = R$ and the ellipse (50) at its points of intersection ζ_k $(k = 0,1,2)$, where ζ_k are taken from (29), we shall be convinced that the circle touches the ellipse in the following three cases only: for $t \in (1,+\infty)$ in the positive root $\zeta_0 = r(t) \in (2,+\infty)$ $(\alpha(t) = 0)$, for $t \in (-\infty,-1)$ in the negative root $\zeta_1 = -r(t) \in (-\infty,-2)$ $(\alpha(t) = \pi)$, and for $t \in (-i\infty,0) \cup (0,+i\infty)$ in the pure imaginary root $\zeta_2 = -i\rho(t) \in (+i\infty,0) \cup (0,-i\infty)$ $(\alpha(t) = \pi/2)$. In each of the three cases the other two points of intersection correspond to the complex roots.

In each of the cases $t \in (1,+\infty)$ and $t \in (-\infty,-1)$ the roots ζ_k $(k = 0,1,2)$ are also the points of intersection of the circle $|\zeta - \zeta^*| = R$, the ellipse (50) and their locus: the left branch H_1 of the hyperbola (53) in [1] and the part of its real axis $(2,+\infty)$ and the right branch H_0 and the part $(-\infty,-2)$, respectively. As homofocal curves of second degree, the intersection between the ellipses (50) and the hyperbola (53) in [1] is orthogonal. By means of the first transformation from (2), we find that in the z-plane the hyperbola (53) in [1] has a central equation $(p = p_1+ip_2 \neq 0;\ z = x+iy)$

$$(56)\qquad (\,|p| - 2p_1)x^2 - 4p_2xy + (\,|p| + 2p_1)y^2 - |p|^2 = 0$$

with vertices $\pm\,{}_0\sqrt{-p/3}$ and the corresponding foci $\pm 2\,{}_0\sqrt{-p/3}$. The roots (29) of the cubic equation (1) lie on the hyperbola (56) and its real axis if the coefficients $p \neq 0$ and q are such that the parameter t is real (according to Theorem 4, the roots for $t \in [-1,1]$ lie on the segment $z \in [-2\,{}_0\sqrt{-p/3}, 2\,{}_0\sqrt{-p/3}]$, determined by the foci). In particular, for real coefficients $p>0$ and q $(-\infty<t<+\infty)$ the roots (29) lie on the hyperbola (and its imaginary axis) with canonical equation

$$(57)\qquad \frac{x^2}{|p|/3} - \frac{y^2}{|p|} = 1.$$

If $t \in (-i\infty,+i\infty)$, then we find the locus of the roots of the cubic equation (3) by the method by which we obtained the hyperbola (53) in [1]. Setting $t = iu$, $u \in (-\infty,+\infty)$, let $\zeta = iv$, $v \in (-\infty,+\infty)$ be the pure imaginary root of this equation. Separating this root, we obtain the other two roots $\zeta = \pm\xi + i\eta = \pm\sqrt{3(v^2+4)}\,/2 - iv/2$, whence we find the hyperbola

$$(58)\qquad \frac{\xi^2}{3} - \eta^2 = 1$$

with vertices $\pm\sqrt{3}$ and foci ± 2. Hence, in the case $t \in (-i\infty,0) \cup (0,+i\infty)$ the roots ζ_k $(k = 0,1,2)$ are also points of intersection of the locus found (the hyperbola (58) and its imaginary axis) with the circle $|\zeta - \zeta^*| = R$ and the homofocal ellipse (50), the intersection with which is orthogonal. By means of the first transformation from (2), we find that in the z-plane the hyperbola (58) has a central equation

$$(59) \qquad (|p| + 2p_1)x^2 + 4p_2xy + (|p| - 2p_1)y^2 + |p|^2 = 0$$

with vertices $\pm{}_0\sqrt{-p}$ and corresponding foci $\pm 2\,{}_0\sqrt{-p/3}$. The roots (29) of the cubic equation (1) lie on the hyperbola (59) and its imaginary axis if the coefficients $p \neq 0$ and q are such that the parameter t is a pure imaginary number. In particular, for real coefficients $p < 0$ and q $(t \in (-i\infty, +i\infty))$ the roots (29) lie on the hyperbola (and its real axis) with canonical equation

$$(60) \qquad \frac{y^2}{p} - \frac{x^2}{p/3} = 1.$$

In the general case, the positions of the roots ζ_k $(k = 0, 1, 2)$ on the ellipse (50) are determined by means of the circle $|\zeta - \zeta^*| = R$ in view of which it is appropriate to call this circle d e t e r m i n i n g. For finite t its center is always different from the origin. From the relations $\alpha(-t) = \pi - \alpha(t)$, $r(-t) = r(t)$ and $\rho(-t) = -\rho(t)$ $(a(-t) = 1/a(t))$ it follows that for values t and $-t$, the centers of two such determining circles are symmetric with respect to the origin and their radii are equal to each other. The fourth point of intersection $\zeta_3 = \xi_3 + i\eta_3$ of the determining circle $|\zeta - \zeta^*| = R$ $(\zeta = \xi + i\eta)$ and the ellipse (50) in the general case is obtained by the joint solution of their equations. Thus, we obtain for ξ and η the corresponding equations of fourth degree which must be satisfied by the real and imaginary parts of the roots ζ_k $(k = 0, 1, 2)$, taken from (29), and of the point ζ_3. In these equations the coefficients of ξ^4 and ξ^3 are 4 and $-4r(t)\cos\alpha(t)$, respectively, and of η^4 and η^3 are 4 and $4\rho(t)\sin\alpha(t)$, respectively . Taking into consideration that $\zeta_0 + \zeta_1 + \zeta_2 = 0$, we obtain

$$(61) \qquad \zeta_3 = r(t)\cos\alpha(t) - i\rho(t)\sin\alpha(t) \qquad (t \notin [-1, 1]).$$

By means of (43) the obtained formula can be immediately expressed by the real and imaginary parts of the roots $\zeta_k = \xi_k + i\eta_k$ $(k = 0, 1, 2)$. From (61) and (47) it follows that the product $\zeta^*\zeta_3$ lies on the permanent straight line $\Re\zeta = 1$.

For the considered values of t the point ζ_3 is always different from the origin. Transformating it by the first transformation in (2), we obtain the point $z_3 = \zeta_3\,{}_0\sqrt{-p/3}$. The position of z_3 on the determining circle (47) in the z-plane with respect to the roots (29) is determined by the values of the cross-ratios

$$(62) \quad (z_0, z_1, z_2, z_3) = \frac{z_2 - z_0}{z_2 - z_1} : \frac{z_3 - z_0}{z_3 - z_1} = \frac{|z_2 - z_0|^2 \sin\frac{2\alpha(t)+\pi}{3}}{|z_2 - z_1|^2 \sin\frac{2\alpha(t)}{3}} \qquad (t \notin [-1, 1]),$$

(63) $$(z_2, z_0, z_1, z_3) = \frac{z_1 - z_2}{z_1 - z_0} : \frac{z_3 - z_2}{z_3 - z_0} = \frac{|z_1 - z_2|^2 \sin \frac{2\alpha(t)}{3}}{|z_1 - z_0|^2 \sin \frac{2\alpha(t)-\pi}{3}} \quad (t \notin [-1,1])$$

evidently invariant to the cross-ratios $(\zeta_0, \zeta_1, \zeta_2, \zeta_3)$ and $(\zeta_2, \zeta_0, \zeta_1, \zeta_3)$, respectively. For $t \in (1, +\infty)$ $(\alpha(t) = 0)$, $t \in (-\infty, -1)$ $(\alpha(t) = \pi)$ and $t \in (-i\infty, 0) \cup (0, +i\infty)$ $(\alpha(t) = \pi/2)$ the values (62) and (63) are ∞, 0, 1 and 0, 1, ∞ or for these t we shall have $z_3 = z_0$, $z_3 = z_1$ and $z_3 = z_2$, respectively. These unique cases of coincidence corresponding to tangency of the determining circle (47) and the ellipse (48) are already obtained by us above in another way. Except these special values of t, for all the remaining finite values of $t \notin [-1, 1]$ (hence $0 < \alpha(t) < \pi$, $\alpha(t) \neq \pi/2$) we have the following position of the points z_k $(k = 0, 1, 2, 3)$ on the determining circle (47): the cross-ratio (62) is positive and different from 1 or the points z_3 and z_2 are situated on one side of the straight line z_0z_1; the cross-ratio (63) is positive and different from 1 for $\Re t < 0$ $(\pi/2 < \alpha(t) < \pi)$ or the points z_3 and z_1 lie on one side of the straight line z_2z_0 (z_3 is between z_1 and z_2) and negative for $\Re t > 0$ $(0 < \alpha(t) < \pi/2)$ or the points z_3 and z_1 lie on both sides of the straight line z_2z_0 (z_3 is between z_2 and z_0).

Hence, for $t \notin [-1, 1]$ each cubic equation (1) with roots z_k $(k = 0, 1, 2)$ has an associated to it corresponding algebraic equation of fourth degree with roots z_k $(k = 0, 1, 2, 3)$, lying on the determining circle (47). Conversely, separating from the cubic equation (1) one or two roots, we obtain corresponding algebraic equations of second and first degree. Thus for $t \notin [-1, 1]$ we obtain the class of all algebraic equations from the first to the fourth degree, the roots of which are points of intersection of the two curves of second degree (47) and (48).

Finally, we shall indicate new metric relations for the roots z_k $(k = 0, 1, 2)$ of the cubic equation (1). By the square of the moduli of the roots (29), we obtain the sharp estimates

(64) $$\sqrt{\rho^2(t)+1}\,\sqrt{|p|/3} \le |z_{0,1}| \le r(t)\,\sqrt{|p|/3} \qquad (\forall t),$$

where the equality is attatined: on the left-hand side for $t \le -1$ by $|z_0|$ and for $t \ge 1$ by $|z_1|$, and on the right-hand side for $t \ge 1$ by $|z_0|$ and for $t \le -1$ by $|z_1|$, only, and

(65) $$|\rho(t)|\,\sqrt{|p|/3} \le |z_2| \le \sqrt{\rho^2(t)+1}\,\sqrt{|p|/3} \qquad (\forall t),$$

where we have equality: on the left-hand side for $t \in (-i\infty, +i\infty)$, and on the right-hand side for $t \le -1$ and $t \ge 1$ only.

Comparing the differences of the square of the moduli of the differences (45) and taking into consideration the first formula from (2), we find sharp inequalities between the distances of the roots (29):

$$(66)\qquad |z_0 - z_1| \overset{t\leq -1}{\geq} |z_2 - z_1|, \quad |z_1 - z_0| \overset{t\geq 1}{\geq} |z_2 - z_0|, \qquad (\forall t).$$

From here, in particular, for $t \notin [-1,1]$ it follows that the largest side of the non-degenerate triangle $z_0z_1z_2$ by the roots (29) of the cubic equation (1) is the side z_0z_1. By the cosine theorem it follows that the angle opposite the side z_0z_1 can be acute, right, or obtuse.

For the roots (29) of the cubic equation (1) we have the equality

$$(67)\qquad |z_0|^2 + |z_1|^2 + |z_2|^2 = |p|\,(r^2(t)-2) \qquad (\forall t)$$

and the following from (45) and from the first formula in (2) equality

$$(68)\qquad |z_0 - z_1|^2 + |z_1 - z_2|^2 + |z_2 - z_0|^2 = 3|p|\,(r^2(t)-2) \qquad (\forall t),$$

the right-hand sides of which, in particular, for $t \in [-1,1]$ $(r(t) = 2)$ are equal to $2|p|$ and $6|p|$, respectively.

The area of the triangle $\zeta_0\zeta_1\zeta_2$ from the roots (29) of the cubic equation (3) is calculated by means of the corresponding determinant, constructed by their real and imaginary parts. Multiplying this area by the square of the modulus of the homothety expressed by the first formula in (2), we obtain the area S of the triangle $z_0z_1z_2$ from the roots (29) of the cubic equation (1):

$$(69)\qquad S = \frac{|p|\sqrt{3}}{4}\, r(t)\,|\rho(t)| \qquad (t \notin [-1,1]).$$

The formula we found is valid for $t \in [-1,1]$ $(\rho(t) = 0)$, as well, when the triangle $z_0z_1z_2$ is degenerated in the segment, considering its area equal to zero.

Our geometry in item 4 is different from Fell's geometry in [4].

5. Tabulation of the roots of the irreducible cubic equation

We shall show how this can be made with the help of computers. For $p<0$ and $-1<t<1$ the cubic equation (1) is irreducible. Considering more generally $-1 \leq t \leq 1$, let us write its three real roots in the form

$$(70)\qquad z_k = f_k(t)\,{}_+\sqrt{-p/3} \qquad \left(k = 0,1,2;\, t = -\frac{q}{2({}_+\sqrt{-p/3})^3};\, p<0\right),$$

where $\zeta_k = f_k(t)$ $(k = 0,1,2)$ are the three real roots of the irreducible cubic equation (3). In consequence from (47) in [1], in (70) we can restrict ourselves to $0 \le t \le 1$ $(-2(\,_+\sqrt{-p/3})^3 \le q \le 0)$. By means of the function (4), as we have done in the beginning of the item 3.2 in [1], we establish that for $0 \le t \le 1$ the functions $f_{0,1}(t)$ increase from $f_0(0) = \sqrt{3}$ to $f_0(1) = 2$ and from $f_1(0) = -\sqrt{3}$ to $f_1(1) = -1$, respectively, and the function $f_2(t)$ decreases from $f_2(0) = 0$ to $f_2(1) = -1$. Hence, the functions $\zeta_k = f_k(t)$ $(k = 0,1,2)$, being bounded for $0 \le t \le 1$, can be tabulated by any of the theorems proved for their representation. The formulas (68) from Theorem 6 in [1] prove to be the most convenient for this purpose. For $0 \le t \le 1$ the hypergeometric functions in (68) in [1] can be replaced by the corresponding quickly convergent hypergeometric series (65) and (71) in [1] since $0 \le (1-t)/2 \le 1/2$. Let us denote the coefficients in these series by

$$(71) \qquad a_m = \frac{(\frac{-1}{3})_m (\frac{1}{3})_m}{(\frac{1}{2})_m\, m!}, \quad b_m = \frac{(\frac{1}{6})_m (\frac{5}{6})_m}{(\frac{3}{2})_m\, m!}, \qquad (m = 0,1,2,\ldots).$$

They tend monotonously to zero so that the ratios $|a_{m+1}/a_m|$ and b_{m+1}/b_m monotonously increase and tend to 1. This circumstance yields a possibility of finding easily good estimates for the errors

$$(72) \qquad \begin{cases} R_{m+1}\left(\frac{-1}{3},\frac{1}{3};\frac{1}{2};\frac{1-t}{2}\right) = \left|\sum_{\nu=m+1}^{\infty} a_\nu \left(\frac{1-t}{2}\right)^\nu\right| < -\frac{2a_m}{1+t}\left(\frac{1-t}{2}\right)^{m+1}, \\ R_{m+1}\left(\frac{1}{6},\frac{5}{6};\frac{3}{2};\frac{1-t}{2}\right) = \sum_{\nu=m+1}^{\infty} b_\nu \left(\frac{1-t}{2}\right)^\nu < \frac{2b_m}{1+t}\left(\frac{1-t}{2}\right)^{m+1} \\ (m=0,1,2,\ldots;0 \le t<1), \end{cases}$$

where the maxima on the right-hand sides with respect to t are attained for $t = 0$ and are equal to $-a_m/2^m$ and $b_m/2^m$ $(m = 0,1,2,\ldots)$, respectively. Taking in view the factors in front of the series in (68) in [1, Theorem 6], the first maximum is multiplied by 2, and the second one is replaced by a smaller one. Of course, the greatest accuracy is obtained if for a given t $(0<t<1)$ the errors are calculated from (72) and the absolute values of the factors in (68) in [1] are taken into consideration.

Hence, there emerges the possibility of constituting tables for the functions (68) in [1, Theorem 6] for $0 \le t< 1$. They will give the roots $\zeta_k = f_k(t)$ $(k = 0,1,2)$ of the irreducible cubic equation (3) so many correct decimal digits as are taken in the increment of t. If for an increment of t, 0.0001 or 0.00001, *etc.*, are taken, we shall have tables with 4 or 5 decimal digits, *etc.* These tables will give the roots (70) of the irreducible cubic equation (1), as well, an exactness to the factor $_+\sqrt{-p/3}$. This result is of great significance for practice and completely satisfies it. For greater accuracy, the table values multiplied by $+\sqrt{-p/3}$ can be used as initial approximations for iterative processes.

In the other two cases, when the cubic equation (1) is with real coefficients $p \neq 0$ and q, the functions representing the roots of the reduced cubic equation (3) are unbounded. In the case $p>0$ or $t \in (-i\infty, +i\infty)$ the roots are represented again by the formula (70), but for values of p and t considered now and for the principal values of the radical. It follows from (47) in [1] that we can restrict ourselves for $t \in [0, +i\infty)$ $(-\infty < q \leq 0)$. For any $t \in [0, +i\infty)$ the functions $f_k(t)$ $(k = 0, 1, 2)$ in (70) can be represented by the hypergeometric series according to that of the Theorems 3, 5, 6, or 7 in [1] for which the moduli of the arguments of the corresponding hypergeometric functions are not larger than 1. If the moduli of the arguments are near 1, an improvement of the convergence of the series is attained by means of the basic functional relations of the Gauss hypergeometric function [5] as we already noted in the end of item 3.4 in [1]. In the case $p < 0$ or $t \leq -1$ and $t \geq 1$ we can restrict ourselves for $t \geq 1$ $(-\infty < q \leq -2(+\sqrt{-p/3})^3)$ according to (47) in [1]. The roots are represented by replacing $f_k^+(t)$ in (70) by ζ_k^+ $(k = 0, 1, 2)$. Herewith the results in [1], namely: Theorem 4, Theorem 6 (the corresponding formulas from (69)), or Theorem 7 (the formulas (76) for $\arg t = 0$) are applied analogously. Hence, in these two cases, obtaining corresponding quickly convergent hypergeometric series, we can constitute tables for the roots in such local sections of t which are encountered and have significance in practice.

From the results obtained in this item it follows that the "irreducible case" of the cubic equation (1) or (3) is represented as the most favorable and suitable for tabulating the roots as well.

References

1. P. G. Todorov. New analytic solution of the Trinomial Algebraic Equation $z^n + pz + q = 0$ by Means of the Goursat Hypergeometric Function I. *Math. Balk.*, **8**, 1994, No. 4, 375–396.
2. G. Belardinelli. Functions hypergéométriques de plusieurs variables et résolution analytique des équations algébriques générales. *Mémorial Sci. Math.*, **145** Gauthier-Villars, Paris, 1960.
3. M. Marden. Geometry of polynomials, Second edition. *Amer. Math. Soc. Surveys & Monographs*, No. 3, Providence, Rhode Island, 1989.
4. H. Fell. The geometry of zeros of trinomial equations. *Rendiconti del Circolo Matematico di Palermo, Serie II*, **XXIX**, 1980, 303–336.

5. A. Erdélyi. Higher Transcendental Functions, vol. I. McGraw-Hill Book Co., Inc., New York, 1953.

Department of Mathematics,
Paisii Hilendarski University,
4000 Plovdiv
BULGARIA

Received 05.04.1993