Solution Formulas for Cubic Equations Without or With Constraints

Ting Zhao, Dongming Wang, Hoon Hong

Solution Formulas for Cubic Equations Without or With Constraints

Ting Zhao [a,b,*], Dongming Wang [b], Hoon Hong [c]

[a] *LMIB – SKLSDE – School of Mathematics and Systems Science, Beihang University, Beijing 100191, China*

[b] *Laboratoire d'Informatique de Paris 6, Université Pierre et Marie Curie – CNRS, 4 place Jussieu – BP 169, 75252 Paris cedex 05, France*

[c] *Department of Mathematics, North Carolina State University, Box 8205, Raleigh, NC 27695-8205, USA*

Abstract

We present a convention (for square/cubic root) which provides correct interpretations of Lagrange's formula for all cubic polynomial equations with real coefficients. Using this convention, we also present a real solution formulas for the general cubic equation with real coefficients under equality and inequality constraints.

Key words: cubic polynomial, solution formula, root convention, constraint

---

1. Introduction

Let $f(x) = x^3 + a_2x^2 + a_1x + a_0 \in \mathbb{C}[x]$, where $\mathbb{C}$ denotes the field of complex numbers. Lagrange (Lagrange, 1770; Smith, 2003) gave the following formula for the three solutions

* Corresponding author.

Email addresses: `zhaoting522@163.com` (Ting Zhao), `dongming.wang@lip6.fr` (Dongming Wang), `hong@math.ncsu.edu` (Hoon Hong).

Preprint submitted to Journal of Symbolic Computation 6 February 2011

u_1, u_2, u_3 of the equation $f(x) = 0$: [1]

$$u_1 = (-a_2 + \omega^1 c_1 + \omega^2 c_2)/3, \quad c_1 = \sqrt[3]{(p_2 + 3\,s)/2}, \quad s = \sqrt[2]{-3\,p_1},$$

$$u_2 = (-a_2 + \omega^0 c_1 + \omega^0 c_2)/3, \quad c_2 = \sqrt[3]{(p_2 - 3\,s)/2},$$

$$u_3 = (-a_2 + \omega^2 c_1 + \omega^1 c_2)/3,$$

$$\begin{aligned}
\omega &= \mathrm{e}^{i\frac{2\pi}{3}} = -\tfrac{1}{2} + \tfrac{\sqrt{3}}{2} i,\\
p_1 &= a_2^2 a_1^2 + 18\,a_2 a_1 a_0 - 4\,a_1^3 - 27\,a_0^2 - 4\,a_2^3 a_0,\\
p_2 &= 9\,a_2 a_1 - 27\,a_0 - 2\,a_2^3.
\end{aligned}$$

Note that the formula, as usually stated, is a bit ambiguous since there are two possible values of s, three possible values of c_1, and three possible values of c_2, depending which square/cubic roots one takes. Hence there are all together $2 \times 3 \times 3 = 18$ possible interpretations of the above formula. It is well known that some interpretations are correct (yielding the solutions), but the others are not.

How to choose a correct interpretation? The usual answer, in the literature, is to choose an interpretation satisfying the condition

$$c_1 c_2 = a_2^2 - 3a_1.$$

Note that the above condition depends on the polynomial f. So we question whether there is a *uniform* condition, i.e., a condition that is *independent* of the polynomial f. The question essentially amounts to whether there is a *convention* for choosing square root and cubic root that will yield correct interpretations for *all* f. We ask the question because it seems to be natural and interesting on its own. We are also motivated by the need of such a convention in geometric constraint solving (Wang, 2004; Hong et al., 2006), where it is very desirable to have a uniform way (independent of f) to choose a correct interpretation.

It is easy to verify that the "standard" convention

$$\arg \sqrt[2]{x} = \frac{1}{2} \arg x, \quad \arg \sqrt[3]{x} = \frac{1}{3} \arg x$$

is *not* always correct. For example, the Lagrange formula under the standard convention on

$$f = x^3 - 2\,x^2 + x = (x-1)^2 x$$

yields the incorrect solutions: $\frac{1}{2} - \frac{\sqrt{3}}{6} i, 1 + \frac{\sqrt{3}}{3} i, \frac{1}{2} - \frac{\sqrt{3}}{6} i$.

[1] The given formula (usually attributed to Lagrange and based on his idea of resolvent) is inspired by but different from the well known formula due to Ferro (communicated by Cardano) (Guilbeau, 1930; Gardano, 1993). Ferro–Cardano's formula involves division. Thus it may encounter a numerically unstable case (i.e., near "0/0" case), when both the numerator and the denominator are close to zero. Lagrange's formula does not require division and thus avoids the "0/0" case. In various applications, such as geometric constraint solving, one needs to solve equations with gradually changing coefficients, for which Ferro–Cardano's formula can encounter near "0/0", resulting in significant numerical errors. Therefore, Lagrange's formula is better for such applications.

Of course there are infinitely many other (non-standard) conventions. However, we do not yet know if there exists a non-standard but correct convention. Nevertheless, in most applications the polynomials have only real coefficients . So we ask instead whether there is a convention that always yields correct solutions if we restrict the coefficients of the polynomials to *real* numbers. The answer is *Yes.* [2]

In the following section (Section 2) we will present the non-standard convention (which we will call "real" convention) that yields correct solutions for all cubic polynomials with real coefficients. In Section 3, we will prove its correctness. In Section 4, using the real convention, we will present real solution formulas for the general real-coefficient cubic equation under equality and inequality constraints. We will prove its correctness in Section 5. Constraints naturally arise in applications such as geometric constraint solving (Wang, 2004; Hong et al., 2006).

2. Real convention

We discovered a correct convention for all cubic equations with real coefficients. The new convention is described in the following definition, under the name of *real convention.*

Definition 1 (Real Convention)**.** The *real convention* (Figure 1) chooses the square root $\sqrt[2]{x}$ and cubic root $\sqrt[3]{x}$ of x so that

$$\arg \sqrt[2]{x} = \frac{1}{2}\arg x,$$

$$\arg \sqrt[3]{x} = \begin{cases} \frac{1}{3}\arg x - \frac{2}{3}\pi & \text{if} \quad -\pi < \arg x < -\frac{\pi}{2}, \\ +\frac{\pi}{2} & \text{if} \quad -\frac{\pi}{2} = \arg x, \\ \frac{1}{3}\arg x & \text{if} \quad -\frac{\pi}{2} < \arg x < +\frac{\pi}{2}, \\ -\frac{\pi}{2} & \text{if} \quad +\frac{\pi}{2} = \arg x, \\ \frac{1}{3}\arg x + \frac{2}{3}\pi & \text{if} \quad +\frac{\pi}{2} < \arg x \le +\pi. \end{cases}$$

Remark 2. The real convention for the square root is the same as the standard one, but for the cubic root it is quite different from the standard one.

[2] One might wonder whether there is any relationship between our question and Bombelli's (O'Connor and Rovertson, 2010; Bortolotti, 1966), since both address the issue of "complex/real numbers" in the context of solving cubic equations. They are completely different questions. Bombelli asked how to deal with the cases where intermediate results involve square root of negative numbers. He developed a theory of complex numbers by analogy with known rules for real numbers and demonstrated that real roots can be obtained even though some intermediate results are non-real numbers. Our question is to find a "uniform convention" (for square/cubic roots) that does *not* depend on the coefficients of the polynomials and that provides correct interpretations of Lagrange's formula for all cubic polynomial equations with real coefficients.

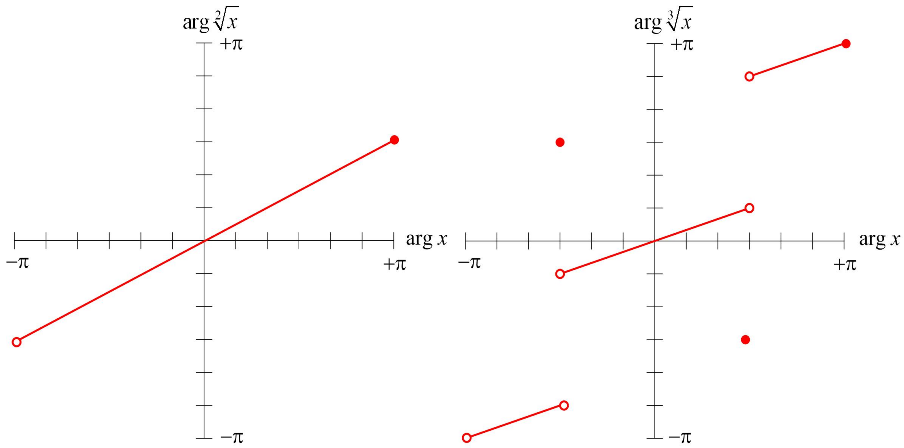

Fig. 1. Real Convention for the Square and Cubic Root

Theorem 3. The Lagrange formula under the real convention yields the correct solutions for all cubic polynomials with real coefficients, and the solution u_2 is always real.

Proof. Will be given in the next section. □

Example 4. We use the example

$$f = x^3 - 2\,x^2 + x = (x-1)^2 x$$

from the introduction to verify the correctness of the real convention. Direct calculations, following the real convention, yield

$$\begin{aligned}
p_1 &= a_2^2 a_1^2 + 18\,a_2 a_1 a_0 - 4\,a_1^3 - 27\,a_0^2 - 4\,a_2^3 a_0 = 0,\\
p_2 &= 9\,a_2 a_1 - 27\,a_0 - 2\,a_2^3 = -2,\\
s &= \sqrt[2]{-3\,p_1} = \sqrt[2]{0} = 0,\\
c_1 &= \sqrt[3]{(p_2 + 3\,s)/2} = \sqrt[3]{-1} = \sqrt[3]{\mathrm{e}^{i\pi}} = \mathrm{e}^{i\pi} = -1,\\
c_2 &= \sqrt[3]{(p_2 - 3\,s)/2} = \sqrt[3]{-1} = \sqrt[3]{\mathrm{e}^{i\pi}} = \mathrm{e}^{i\pi} = -1,\\
u_1 &= (-a_2 + \omega^1 c_1 + \omega^2 c_2)/3 = 1,\\
u_2 &= (-a_2 + \omega^0 c_1 + \omega^0 c_2)/3 = 0,\\
u_3 &= (-a_2 + \omega^2 c_1 + \omega^1 c_2)/3 = 1.
\end{aligned}$$

Clearly, u_1, u_2, u_3 are the three solutions of $f = 0$.

Example 5. Consider another polynomial

$$f = x^3 + x = x(x+i)(x-i).$$

Direct calculations, following the real convention, yield

$$\begin{aligned}
p_1 &= -4,\ p_2 = 0,\ s = 2\sqrt[2]{3}, \\
c_1 &= \sqrt[2]{3},\ c_2 = -\sqrt[2]{3}, \\
u_1 &= (-a_2 + \omega^1 c_1 + \omega^2 c_2)/3 = i, \\
u_2 &= (-a_2 + \omega^0 c_1 + \omega^0 c_2)/3 = 0, \\
u_3 &= (-a_2 + \omega^2 c_1 + \omega^1 c_2)/3 = -i.
\end{aligned}$$

Clearly, u_1, u_2, u_3 are the three solutions of $f = 0$ and u_2 is real.

3. Proof of the correctness of the real convention

In this section, we prove Theorem 3 stated in the previous section. Let f be an arbitrary (monic) cubic polynomial. Let r_1, r_2, r_3 be the three (complex) solutions of $f = 0$. Using the well-known relations

$$\begin{aligned}
a_2 &= -r_1 - r_2 - r_3, \\
a_1 &= r_1 r_2 + r_1 r_3 + r_2 r_3, \\
a_0 &= -r_1 r_2 r_3,
\end{aligned}$$

we can rewrite p_1 and p_2 as

$$\begin{aligned}
p_1 &= (r_1 - r_2)^2 (r_1 - r_3)^2 (r_2 - r_3)^2, \\
p_2 &= (2r_1 - r_2 - r_3)(2r_2 - r_1 - r_3)(2r_3 - r_1 - r_2).
\end{aligned}$$

It is easy to verify that the signs of p_1 and p_2 determine the "configuration" of the solutions r_1, r_2 and r_3, as shown in Figure 2. We have also indexed the solutions so that we can refer to them later on. Note that the indexing for the bottom-middle configuration is peculiar (causing solutions jump discontinuously) but it is essential.

The proof proceeds by rewriting, in terms of the solutions, the expressions for s, c_1, c_2 and u_1, u_2, u_3 in Lagrange's formula, taking radicals according to the real convention. It is split into the following several lemmas.

Lemma 6. $s = i\sqrt{3}\,(r_1 - r_2)(r_1 - r_3)(r_2 - r_3)$.

Proof. Let $q = -3\,p_1$. Then we obviously have

$$q = \left[i\sqrt{3}\,(r_1 - r_2)(r_1 - r_3)(r_2 - r_3)\right]^2.$$

Hence $\sqrt{q}$ is one of the following:

$$\begin{aligned}
q_1 &= +i\sqrt{3}\,(r_1 - r_2)(r_1 - r_3)(r_2 - r_3), \\
q_2 &= -i\sqrt{3}\,(r_1 - r_2)(r_1 - r_3)(r_2 - r_3).
\end{aligned}$$

We proceed to show that $s = q_1$ in every configuration of the solutions.

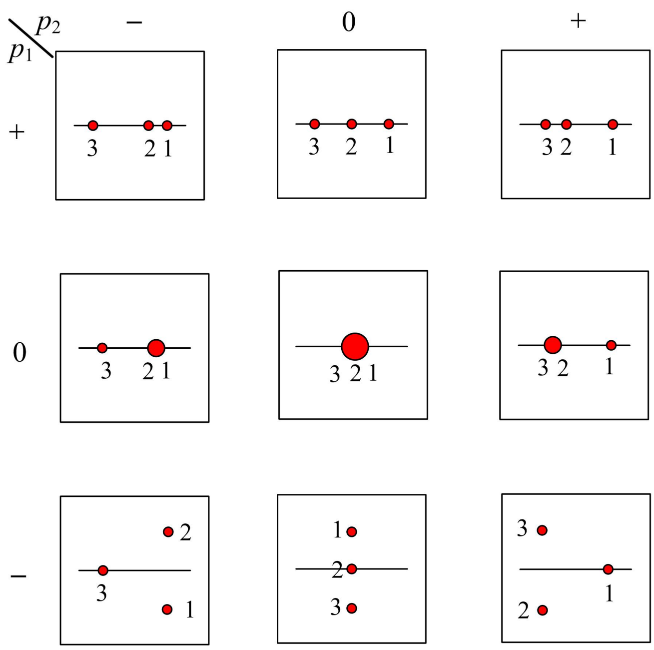

Fig. 2. Solution Indexing. Each rectangle denotes a complex plane, in which the horizontal line is the real axis with left-to-right direction. A small disk stands for a simple solution, a bigger disk for a double solution, and the biggest disk for a triple solution.

(1) $p_1 > 0$. In this case, $f = 0$ has three real solutions indexed as $r_3 < r_2 < r_1$. Note that

$$\arg q_1 = +\frac{\pi}{2}, \quad \arg q_2 = -\frac{\pi}{2}.$$

Hence $\sqrt{q} = q_1$. Thus $s = q_1$.

(2) $p_1 = 0$. In this case, $f = 0$ has a multiple solution. It follows that $q_1 = q_2 = 0$ and

$$\arg q_1 = 0, \quad \arg q_2 = 0.$$

Hence $\sqrt{q} = q_1$. Thus $s = q_1$.

(3) $p_1 < 0$ and $p_2 > 0$. In this case, $f = 0$ has a real solution r_1 and a pair of complex conjugates $r_3 = \alpha + i\beta$ and $r_2 = \alpha - i\beta$ such that $r_1 > \alpha$ and $\beta > 0$. Simple calculation shows that

$$q_1 = +2\sqrt{3}\,\beta\left[(r_1 - \alpha)^2 + \beta^2\right] > 0,$$
$$q_2 = -2\sqrt{3}\,\beta\left[(r_1 - \alpha)^2 + \beta^2\right] < 0.$$

Then

$$\arg q_1 = 0, \quad \arg q_2 = \pi.$$

Hence $\sqrt{q} = q_1$. Thus $s = q_1$.

(4) $p_1 < 0$ and $p_2 = 0$. In this case, $f = 0$ has a real solution r_2 and a pair of complex conjugates $r_1 = \alpha + i\beta$ and $r_3 = \alpha - i\beta$ such that $r_2 = \alpha$ and $\beta > 0$. Simple calculation shows that

$$q_1 = +2\sqrt{3}\,\beta^3 > 0, \quad q_2 = -2\sqrt{3}\,\beta^3 < 0.$$

Then

$$\arg q_1 = 0, \quad \arg q_2 = \pi.$$

Hence $\sqrt{q} = q_1$. Thus $s = q_1$.

(5) $p_1 < 0$ and $p_2 < 0$. In this case, $f = 0$ has a real solution r_3 and a pair of complex conjugates $r_2 = \alpha + i\beta$ and $r_1 = \alpha - i\beta$ such that $r_3 < \alpha$ and $\beta > 0$. Simple calculation shows that

$$\begin{aligned} q_1 &= +2\sqrt{3}\,\beta\left[(r_3-\alpha)^2+\beta^2\right] > 0, \\ q_2 &= -2\sqrt{3}\,\beta\left[(r_3-\alpha)^2+\beta^2\right] < 0. \end{aligned}$$

Then

$$\arg q_1 = 0, \quad \arg q_2 = \pi.$$

Hence $\sqrt{q} = q_1$. Thus $s = q_1$.

□

Lemma 7. At least one of the followings is true.

$$\begin{aligned} c_1 &= \omega^0 r_1 + \omega^1 r_2 + \omega^2 r_3 \quad \wedge \quad c_2 = \omega^0 r_1 + \omega^2 r_2 + \omega^1 r_3 \\ c_1 &= \omega^2 r_1 + \omega^0 r_2 + \omega^1 r_3 \quad \wedge \quad c_2 = \omega^1 r_1 + \omega^0 r_2 + \omega^2 r_3 \\ c_1 &= \omega^1 r_1 + \omega^2 r_2 + \omega^0 r_3 \quad \wedge \quad c_2 = \omega^2 r_1 + \omega^1 r_2 + \omega^0 r_3 \end{aligned}$$

Proof. Let $q = (p_2 + 3\,s)/2$ and $q' = (p_2 - 3\,s)/2$. Recalling Lemma 6, substitution and factorization yield

$$q = (\omega^0 r_1 + \omega^1 r_2 + \omega^2 r_3)^3, \quad q' = (\omega^0 r_1 + \omega^2 r_2 + \omega^1 r_3)^3.$$

Hence $\sqrt[3]{q}$ is one of the following:

$$\begin{aligned} q_1 &= \omega^0 r_1 + \omega^1 r_2 + \omega^2 r_3, \\ q_2 &= \omega^2 r_1 + \omega^0 r_2 + \omega^1 r_3, \\ q_3 &= \omega^1 r_1 + \omega^2 r_2 + \omega^0 r_3. \end{aligned}$$

Likewise $\sqrt[3]{q'}$ is one of the following:

$$\begin{aligned} q'_1 &= \omega^0 r_1 + \omega^2 r_2 + \omega^1 r_3, \\ q'_2 &= \omega^1 r_1 + \omega^0 r_2 + \omega^2 r_3, \\ q'_3 &= \omega^2 r_1 + \omega^1 r_2 + \omega^0 r_3. \end{aligned}$$

We can rewrite q_1, q_2, q_3 and q'_1, q'_2, q'_3 as

$$\begin{aligned} q_1 &= \omega^0(r_1-r_2) - \omega^2(r_2-r_3) = \mathrm{e}^{i\frac{+0}{6}\pi}(r_1-r_2) + \mathrm{e}^{i\frac{+2}{6}\pi}(r_2-r_3), \\ q_2 &= \omega^2(r_1-r_2) - \omega^1(r_2-r_3) = \mathrm{e}^{i\frac{-4}{6}\pi}(r_1-r_2) + \mathrm{e}^{i\frac{-2}{6}\pi}(r_2-r_3), \\ q_3 &= \omega^1(r_1-r_2) - \omega^0(r_2-r_3) = \mathrm{e}^{i\frac{+4}{6}\pi}(r_1-r_2) + \mathrm{e}^{i\frac{+6}{6}\pi}(r_2-r_3); \end{aligned}$$

$$\begin{aligned} q'_1 &= \omega^0(r_1-r_2) - \omega^1(r_2-r_3) = \mathrm{e}^{i\frac{+0}{6}\pi}(r_1-r_2) + \mathrm{e}^{i\frac{-2}{6}\pi}(r_2-r_3), \\ q'_2 &= \omega^1(r_1-r_2) - \omega^2(r_2-r_3) = \mathrm{e}^{i\frac{+4}{6}\pi}(r_1-r_2) + \mathrm{e}^{i\frac{+2}{6}\pi}(r_2-r_3), \\ q'_3 &= \omega^2(r_1-r_2) - \omega^0(r_2-r_3) = \mathrm{e}^{i\frac{-4}{6}\pi}(r_1-r_2) + \mathrm{e}^{i\frac{+6}{6}\pi}(r_2-r_3). \end{aligned}$$

Now we prove the lemma for every configuration of the solutions.

(1) $p_1 > 0$ and $p_2 > 0$. In this case, $f = 0$ has three real solutions $r_3 < r_2 < r_1$ and $r_2 - r_3 < r_1 - r_2$. Thus

$$\begin{aligned} \frac{0}{6}\pi < \arg q_1 < \frac{+\frac{0}{6}\pi + \frac{2}{6}\pi}{2} &= +\frac{1}{6}\pi, \\ -\frac{4}{6}\pi < \arg q_2 < \frac{-\frac{4}{6}\pi - \frac{2}{6}\pi}{2} &= -\frac{3}{6}\pi, \\ +\frac{4}{6}\pi < \arg q_3 < \frac{+\frac{4}{6}\pi + \frac{6}{6}\pi}{2} &= +\frac{5}{6}\pi; \end{aligned}$$

$$\begin{aligned} -\frac{1}{6}\pi = \frac{-\frac{2}{6}\pi + \frac{0}{6}\pi}{2} &< \arg q_1' < +\frac{0}{6}\pi, \\ +\frac{3}{6}\pi = \frac{+\frac{4}{6}\pi + \frac{2}{6}\pi}{2} &< \arg q_2' < +\frac{4}{6}\pi, \\ -\frac{5}{6}\pi = \frac{-\frac{4}{6}\pi - \frac{6}{6}\pi}{2} &< \arg q_3' < -\frac{4}{6}\pi. \end{aligned}$$

Since $\operatorname{Re} q = \operatorname{Re} q' = p_2/2$ (where $\operatorname{Re} q$ denotes the real part of q), we have $|\arg q| < \pi/2$, $|\arg q'| < \pi/2$. Therefore $0 \le |\arg \sqrt[3]{q}| < \pi/6$, $0 \le |\arg \sqrt[3]{q'}| < \pi/6$. Hence $\sqrt[3]{q} = q_1$, $\sqrt[3]{q'} = q_1'$. So we have $c_1 = q_1$, $c_2 = q_1'$.

(2) $p_1 > 0$ and $p_2 = 0$. In this case, $f = 0$ has three real solutions $r_3 < r_2 < r_1$ and $r_2 - r_3 = r_1 - r_2$. Thus

$$\begin{aligned} \arg q_1 &= \frac{+\frac{0}{6}\pi + \frac{2}{6}\pi}{2} = +\frac{1}{6}\pi, \\ \arg q_2 &= \frac{-\frac{4}{6}\pi - \frac{2}{6}\pi}{2} = -\frac{3}{6}\pi, \\ \arg q_3 &= \frac{+\frac{4}{6}\pi + \frac{6}{6}\pi}{2} = +\frac{5}{6}\pi; \end{aligned}$$

$$\begin{aligned} \arg q_1' &= \frac{+\frac{0}{6}\pi - \frac{2}{6}\pi}{2} = -\frac{1}{6}\pi, \\ \arg q_2' &= \frac{+\frac{2}{6}\pi + \frac{4}{6}\pi}{2} = +\frac{3}{6}\pi, \\ \arg q_3' &= \frac{-\frac{4}{6}\pi - \frac{6}{6}\pi}{2} = -\frac{5}{6}\pi. \end{aligned}$$

Since $\operatorname{Re} q = \operatorname{Re} q' = p_2/2 = 0$, we have $|\arg q| = \pi/2$, $|\arg q'| = \pi/2$. Therefore $|\arg \sqrt[3]{q}| = \pi/2$, $|\arg \sqrt[3]{q'}| = \pi/2$. Hence $\sqrt[3]{q} = q_2$, $\sqrt[3]{q'} = q_2'$. So we have $c_1 = q_2$, $c_2 = q_2'$.

(3) $p_1 > 0$ and $p_2 < 0$. In this case, $f = 0$ has three real solutions $r_3 < r_2 < r_1$ and $r_1 - r_2 < r_2 - r_3$. Thus

$$\begin{aligned} +\frac{1}{6}\pi = \frac{+\frac{0}{6}\pi + \frac{2}{6}\pi}{2} &< \arg q_1 < +\frac{2}{6}\pi, \\ -\frac{3}{6}\pi = \frac{-\frac{4}{6}\pi - \frac{2}{6}\pi}{2} &< \arg q_2 < -\frac{2}{6}\pi, \\ +\frac{5}{6}\pi = \frac{+\frac{4}{6}\pi + \frac{6}{6}\pi}{2} &< \arg q_3 < +\frac{6}{6}\pi; \end{aligned}$$

$$
\begin{aligned}
-\frac{2}{6}\pi < \arg q_1' &< \frac{+\frac{0}{6}\pi - \frac{2}{6}\pi}{2} = -\frac{1}{6}\pi,\\
+\frac{2}{6}\pi < \arg q_2' &< \frac{+\frac{2}{6}\pi + \frac{4}{6}\pi}{2} = +\frac{3}{6}\pi,\\
-\frac{6}{6}\pi < \arg q_3' &< \frac{-\frac{4}{6}\pi - \frac{6}{6}\pi}{2} = -\frac{5}{6}\pi.
\end{aligned}
$$

Since $\operatorname{Re} q = \operatorname{Re} q' = p_2/2 < 0$, we have $|\arg q\,| > \pi/2$, $|\arg q'\,| > \pi/2$. Therefore $5\,\pi/6 < |\arg \sqrt[3]{q}\,| < \pi$, $5\,\pi/6 < |\arg \sqrt[3]{q'}\,| < \pi$. Hence $\sqrt[3]{q} = q_3$, $\sqrt[3]{q'} = q_3'$. So we have $c_1 = q_3,\ c_2 = q_3'$.

(4) $p_1 = 0$ and $p_2 > 0$. In this case, $f = 0$ has a simple real solution r_1 and a double real solution $r_2 = r_3$ such that $r_3 = r_2 < r_1$. Thus

$$
\begin{aligned}
\arg q_1 &= +\frac{0}{6}\pi, \quad \arg q_2 = -\frac{4}{6}\pi, \quad \arg q_3 = +\frac{4}{6}\pi;\\
\arg q_1' &= +\frac{0}{6}\pi, \quad \arg q_2' = +\frac{4}{6}\pi, \quad \arg q_3' = -\frac{4}{6}\pi.
\end{aligned}
$$

Since $q = q' = p_2/2 > 0$, we have $\arg q = 0$, $\arg q' = 0$. Therefore $\arg \sqrt[3]{q} = 0$, $\arg \sqrt[3]{q'} = 0$. Hence $\sqrt[3]{q} = q_1$, $\sqrt[3]{q'} = q_1'$. So we have $c_1 = q_1,\ c_2 = q_1'$.

(5) $p_1 = 0$ and $p_2 = 0$. In this case, $f = 0$ has a triple real solution $r_3 = r_2 = r_1$. It follows that $q_1 = q_2 = q_3 = 0$, $q_1' = q_2' = q_3' = 0$. Since $q = q' = p_2/2 = 0$, we have $\arg q = 0$, $\arg q' = 0$. Therefore $\arg \sqrt[3]{q} = 0$, $\arg \sqrt[3]{q'} = 0$. Hence we can choose $\sqrt[3]{q} = q_2$, $\sqrt[3]{q'} = q_2'$. So we have $c_1 = q_2,\ c_2 = q_2'$.

(6) $p_1 = 0$ and $p_2 < 0$. In this case, $f = 0$ has a simple real solution r_3 and a double real solution $r_1 = r_2$ such that $r_3 < r_2 = r_1$. Thus

$$
\begin{aligned}
\arg q_1 &= +\frac{2}{6}\pi, \quad \arg q_2 = -\frac{2}{6}\pi, \quad \arg q_3 = +\frac{6}{6}\pi;\\
\arg q_1' &= -\frac{2}{6}\pi, \quad \arg q_2' = +\frac{2}{6}\pi, \quad \arg q_3' = +\frac{6}{6}\pi.
\end{aligned}
$$

Since $q = q' = p_2/2 < 0$, we have $\arg q = \pi$, $\arg q' = \pi$. Therefore $\arg \sqrt[3]{q} = \pi$, $\arg \sqrt[3]{q'} = \pi$. Hence $\sqrt[3]{q} = q_3$, $\sqrt[3]{q'} = q_3'$. So we have $c_1 = q_3,\ c_2 = q_3'$.

(7) $p_1 < 0$ and $p_2 > 0$. In this case, $f = 0$ has a real solution r_1 and a pair of complex conjugates $r_3 = \alpha + i\beta$ and $r_2 = \alpha - i\beta$ such that $r_1 > \alpha$ and $\beta > 0$. Simple calculation gives

$$
\begin{aligned}
q_1 &= \omega^0(r_1 - \alpha + \sqrt{3}\,\beta), \ \ q_1' = \omega^0(r_1 - \alpha - \sqrt{3}\,\beta),\\
q_2 &= \omega^2(r_1 - \alpha + \sqrt{3}\,\beta), \ \ q_2' = \omega^2(r_1 - \alpha - \sqrt{3}\,\beta),\\
q_3 &= \omega^1(r_1 - \alpha + \sqrt{3}\,\beta); \ \ q_3' = \omega^1(r_1 - \alpha - \sqrt{3}\,\beta).
\end{aligned}
$$

Note that

$$
q = (r_1 - \alpha + \sqrt{3}\,\beta)^3 > 0, \quad q' = (r_1 - \alpha - \sqrt{3}\,\beta)^3.
$$

We consider the three subcases.

(a) $r_1 - \alpha - \sqrt{3}\,\beta > 0$. In this case,

$$\arg q_1 = +\frac{0}{3}\pi, \quad \arg q_2 = -\frac{2}{3}\pi, \quad \arg q_3 = +\frac{2}{3}\pi;$$
$$\arg q_1' = +\frac{0}{3}\pi, \quad \arg q_2' = +\frac{2}{3}\pi, \quad \arg q_3' = -\frac{2}{3}\pi.$$

Since $q > 0$, $q' > 0$, we have $\arg q = 0$, $\arg q' = 0$. Therefore $\arg \sqrt[3]{q} = 0$, $\arg \sqrt[3]{q'} = 0$. Hence $\sqrt[3]{q} = q_1$, $\sqrt[3]{q'} = q_1'$.

(b) $r_1 - \alpha - \sqrt{3}\,\beta = 0$. In this case, $q_1 = q_2 = q_3 = 0$, $q_1' = q_2' = q_3' = 0$ and thus

$$\arg q_1 = 0, \quad \arg q_2 = 0, \quad \arg q_3 = 0;$$
$$\arg q_1' = 0, \quad \arg q_2' = 0, \quad \arg q_3' = 0.$$

Since $q = q' = 0$, we have $\arg q = 0$, $\arg q' = 0$. Therefore $\arg \sqrt[3]{q} = 0$, $\arg \sqrt[3]{q'} = 0$. Hence we can choose $\sqrt[3]{q} = q_1$, $\sqrt[3]{q'} = q_1'$.

(c) $r_1 - \alpha - \sqrt{3}\,\beta < 0$. In this case,

$$\arg q_1 = +\frac{0}{3}\pi, \quad \arg q_2 = -\frac{2}{3}\pi, \quad \arg q_3 = +\frac{2}{3}\pi;$$
$$\arg q_1' = +\frac{3}{3}\pi, \quad \arg q_2' = -\frac{1}{3}\pi, \quad \arg q_3' = +\frac{1}{3}\pi.$$

Since $q > 0$, $q' < 0$, we have $\arg q = 0$, $\arg q' = \pi$. Therefore $\arg \sqrt[3]{q} = 0$, $\arg \sqrt[3]{q'} = \pi$. Hence $\sqrt[3]{q} = q_1$, $\sqrt[3]{q'} = q_1'$.

So we have $c_1 = q_1,\ c_2 = q_1'$.

(8) $p_1 < 0$ and $p_2 = 0$. In this case, $f = 0$ has a real solution r_2 and a pair of complex conjugates $r_1 = \alpha + i\beta$ and $r_3 = \alpha - i\beta$ such that $r_2 = \alpha$ and $\beta > 0$. Simple calculation gives

$$q_1 = \omega^1\sqrt{3}\,\beta, \quad q_2 = \omega^0\sqrt{3}\,\beta, \quad q_3 = \omega^2\sqrt{3}\,\beta;$$
$$q_1' = -\omega^2\sqrt{3}\,\beta, \ q_2' = -\omega^0\sqrt{3}\,\beta, \ q_3' = -\omega^1\sqrt{3}\,\beta.$$

Thus

$$\arg q_1 = +\frac{2}{3}\pi, \quad \arg q_2 = +\frac{0}{3}\pi, \quad \arg q_3 = -\frac{2}{3}\pi;$$
$$\arg q_1' = +\frac{1}{3}\pi, \quad \arg q_2' = +\frac{3}{3}\pi, \quad \arg q_3' = -\frac{1}{3}\pi.$$

Since $q = 3s/2 > 0$, $q' = -3s/2 < 0$, we have $\arg q = 0$, $\arg q' = \pi$. Therefore $\arg \sqrt[3]{q} = 0$, $\arg \sqrt[3]{q'} = \pi$. Hence $\sqrt[3]{q} = q_2$, $\sqrt[3]{q'} = q_2'$. So we have $c_1 = q_2,\ c_2 = q_2'$.

(9) $p_1 < 0$ and $p_2 < 0$. In this case, $f = 0$ has a real solution r_3 and a pair of complex conjugates $r_2 = \alpha + i\beta$ and $r_1 = \alpha - i\beta$ such that $r_3 < \alpha$ and $\beta > 0$. Simple calculation gives

$$q_1 = \omega^2(r_3 - \alpha + \sqrt{3}\,\beta), \ q_1' = \omega^2(r_3 - \alpha - \sqrt{3}\,\beta),$$
$$q_2 = \omega^1(r_3 - \alpha + \sqrt{3}\,\beta), \ q_2' = \omega^1(r_3 - \alpha - \sqrt{3}\,\beta),$$
$$q_3 = \omega^0(r_3 - \alpha + \sqrt{3}\,\beta); \ q_3' = \omega^0(r_3 - \alpha - \sqrt{3}\,\beta).$$

Note that

$$q = (r_3 - \alpha + \sqrt{3}\,\beta)^3, \quad q' = (r_3 - \alpha - \sqrt{3}\,\beta)^3 < 0.$$

We consider the three subcases.

(a) $r_3 - \alpha + \sqrt{3}\,\beta > 0$. In this case,

$$\arg q_1 = -\frac{2}{3}\pi, \quad \arg q_2 = +\frac{2}{3}\pi, \quad \arg q_3 = +\frac{0}{3}\pi;$$
$$\arg q'_1 = -\frac{1}{3}\pi, \quad \arg q'_2 = +\frac{1}{3}\pi, \quad \arg q'_3 = +\frac{3}{3}\pi.$$

Since $q > 0,\ q' < 0$, we have $\arg q = 0$, $\arg q' = \pi$. Therefore $\arg \sqrt[3]{q} = 0$, $\arg \sqrt[3]{q'} = \pi$. Hence $\sqrt[3]{q} = q_3,\ \sqrt[3]{q'} = q'_3$.

(b) $r_3 - \alpha + \sqrt{3}\,\beta = 0$. In this case, $q_1 = q_2 = q_3 = 0$, $q'_1 = q'_2 = q'_3 = 0$ and thus

$$\arg q_1 = 0, \qquad \arg q_2 = 0, \qquad \arg q_3 = 0;$$
$$\arg q'_1 = -\frac{1}{3}\pi,\ \arg q'_2 = +\frac{1}{3}\pi,\ \arg q'_3 = +\frac{3}{3}\pi.$$

Since $q = 0,\ q' < 0$, we have $\arg q = 0$, $\arg q' = \pi$. Therefore $\arg \sqrt[3]{q} = 0$, $\arg \sqrt[3]{q'} = \pi$. Hence we can choose $\sqrt[3]{q} = q_3,\ \sqrt[3]{q'} = q'_3$.

(c) $r_3 - \alpha + \sqrt{3}\,\beta < 0$. In this case,

$$\arg q_1 = +\frac{1}{3}\pi, \quad \arg q_2 = -\frac{1}{3}\pi, \quad \arg q_3 = +\frac{3}{3}\pi;$$
$$\arg q'_1 = -\frac{1}{3}\pi, \quad \arg q'_2 = +\frac{1}{3}\pi, \quad \arg q'_3 = +\frac{3}{3}\pi.$$

Since $q < 0,\ q' < 0$, we have $\arg q = \pi$, $\arg q' = \pi$. Therefore $\arg \sqrt[3]{q} = \pi$, $\arg \sqrt[3]{q'} = \pi$. Hence $\sqrt[3]{q} = q_3,\ \sqrt[3]{q'} = q'_3$.

So we have $c_1 = q_3,\ c_2 = q'_3$.

□

Lemma 8. The solution u_2 is always real.

Proof. We use the results and the notations in the proof of Lemma 7.

(1) $p_1 > 0$ and $p_2 > 0$. In this case, we have $c_1 = q_1,\ c_2 = q'_1$. Substituting c_1 and c_2 into u_2 in Lagrange's formula and simplifying the resulting expressions using $\omega^3 = 1$ and $\omega^0 + \omega^1 + \omega^2 = 0$, we see that

$$u_2 = \frac{3\,r_1 + (\omega^0 + \omega^1 + \omega^2)\,r_2 + (\omega^0 + \omega^1 + \omega^2)\,r_3}{3} = r_1.$$

(2) $p_1 > 0$ and $p_2 = 0$. In this case $c_1 = q_2,\ c_2 = q'_2$. Similar calculation yields $u_2 = r_2$.
(3) $p_1 > 0$ and $p_2 < 0$. In this case $c_1 = q_3,\ c_2 = q_3$. Similar calculation yields $u_2 = r_3$.
(4) $p_1 = 0$ and $p_2 > 0$. In this case $c_1 = q_1,\ c_2 = q'_1$. Similar calculation yields $u_2 = r_1$.
(5) $p_1 = 0$ and $p_2 = 0$. In this case $c_1 = q_2,\ c_2 = q'_2$. Similar calculation yields $u_2 = r_2$.
(6) $p_1 = 0$ and $p_2 < 0$. In this case $c_1 = q_3,\ c_2 = q'_3$. Similar calculation yields $u_2 = r_3$.
(7) $p_1 < 0$ and $p_2 > 0$. In this case $c_1 = q_1,\ c_2 = q'_1$. Similar calculation yields $u_2 = r_1$.
(8) $p_1 < 0$ and $p_2 = 0$. In this case $c_1 = q_2,\ c_2 = q'_2$. Similar calculation yields $u_2 = r_2$.
(9) $p_1 < 0$ and $p_2 < 0$. In this case $c_1 = q_3,\ c_2 = q'_3$. Similar calculation yields $u_2 = r_3$.

It is clear that $u_2 = r_1$ when $p_2 > 0$; $u_2 = r_2$ when $p_2 = 0$; $u_2 = r_3$ when $p_2 < 0$. According to the configurations in Figure 2, we see immediately that u_2 is always real. $\square$

Proof of Theorem 3. Recalling Lemma 7, we consider the following three cases.

(1) $c_1 = \omega^0 r_1 + \omega^1 r_2 + \omega^2 r_3 \quad \wedge \quad c_2 = \omega^0 r_1 + \omega^2 r_2 + \omega^1 r_3'$.
Substituting c_1 and c_2 into u_k and simplifying the resulting expressions using $\omega^3 = 1$ and $\omega^0 + \omega^1 + \omega^2 = 0$, we see that

$$\begin{aligned} u_1 &= \frac{3\,r_3 + (\omega^0 + \omega^1 + \omega^2)\,r_1 + (\omega^0 + \omega^1 + \omega^2)\,r_2}{3} = r_3, \\ u_2 &= \frac{3\,r_1 + (\omega^0 + \omega^1 + \omega^2)\,r_2 + (\omega^0 + \omega^1 + \omega^2)\,r_3}{3} = r_1, \\ u_3 &= \frac{3\,r_2 + (\omega^0 + \omega^1 + \omega^2)\,r_1 + (\omega^0 + \omega^1 + \omega^2)\,r_2}{3} = r_2. \end{aligned}$$

(2) $c_1 = \omega^2 r_1 + \omega^0 r_2 + \omega^1 r_3 \quad \wedge \quad c_2 = \omega^1 r_1 + \omega^0 r_2 + \omega^2 r_3'$.
Similar calculation yields $u_1 = r_1,\ u_2 = r_2,\ u_3 = r_3$.

(3) $c_1 = \omega^1 r_1 + \omega^2 r_2 + \omega^0 r_3 \quad \wedge \quad c_2 = \omega^2 r_1 + \omega^1 r_2 + \omega^0 r_3'$.
Similar calculation yields $u_1 = r_2,\ u_2 = r_3,\ u_3 = r_1$.

From Lemma 8, u_2 is always real. $\square$

4. Cubic formula with constraints

In Section 2, we have introduced a correct convention for choosing the square and cubic roots. Using this convention and Lagrange's formula, we present real solution formulas for the general real-coefficient cubic equation under equality and inequality constraints. Constraints naturally arise in applications such as geometric constraint solving (Wang, 2004; Hong et al., 2006). The representations of the real solutions coupled with real constraints are achieved by combining Thom's lemma (Basu et al., 2006, p. 50) and the complex-solution formulas.

Let $\wedge$, $\vee$, $\Rightarrow$, and $\neg$ stand for the logical connectives "and," "or," "imply," and "not" respectively. Denote by $\mathbb{R}$ the field of real numbers and $\mathbb{R}[x]$ the ring of polynomials in x with real coefficients. We have the following result.

Theorem 9. Let $f(x) = x^3 + a_2x^2 + a_1x + a_0 \in \mathbb{R}[x]$ and $\Gamma(x)$ be a formula composed by $\wedge$, $\vee$, $\Rightarrow$, and $\neg$ of polynomial equality and inequality relations in x, the coefficients of $f(x)$, and other parameters. Then for all $x \in \mathbb{R}$,

$$[f(x) = 0 \wedge \Gamma(x)] \iff [x = u_1 \wedge \Gamma_1] \vee [x = u_2 \wedge \Gamma_2] \vee [x = u_3 \wedge \Gamma_3],$$

where

$$\begin{aligned} u_1 &= (-a_2 + \omega^{(1-\sigma)} c_1 + \omega^{(2+\sigma)} c_2)/3, \\ u_2 &= (-a_2 + \omega^{(0-\sigma)} c_1 + \omega^{(0+\sigma)} c_2)/3, \\ u_3 &= (-a_2 + \omega^{(2-\sigma)} c_1 + \omega^{(1+\sigma)} c_2)/3, \\ \sigma &= \operatorname{sign}(p_2), \end{aligned}$$

and

$$\Gamma_j := (\exists\, x \in \mathbb{R})\,[f(x) = 0 \wedge \Gamma(x) \wedge \Phi_j(x)], \quad j = 1, 2, 3,$$

$$\begin{aligned}
\Phi_1(x) &:= [f'(x) > 0 \wedge f''(x) > 0] \vee [f'(x) = 0 \wedge f''(x) \geq 0],\\
\Phi_2(x) &:= [f'(x) \leq 0] \qquad\qquad\qquad\quad \vee [f''(x) = 0],\\
\Phi_3(x) &:= [f'(x) > 0 \wedge f''(x) < 0] \vee [f'(x) = 0 \wedge f''(x) \leq 0].
\end{aligned}$$

Here c_1, c_2, p_2, ω are the same as in Lagrange's formula given in the introduction.

Proof. Will be given in the next section. $\square$

Remark 10. Note that the above formula is slightly different from the Lagrange formula (in the introduction), in that the exponents for ω are adjusted depending on the sign of p_2. This adjustment is essential for the correctness of the theorem.

Remark 11. It turns out (and will be shown in the proof of the theorem) that the three complex solutions of f satisfy

$$\operatorname{Re} u_3 \leq \operatorname{Re} u_2 \leq \operatorname{Re} u_1.$$

Remark 12. The real constraints in the formula are given as three existentially quantified subformulas Γ_j. If needed, one could eliminate the existential quantifier using, e.g., the method based on partial cylindrical algebraic decomposition (Collins and Hong, 1991). However, if $\Gamma(x)$ is restricted to a combination of polynomial equalities and inequalities of degree ≤ 3 in x, one could use the alternative approach of Weispfenning (1994) that provides explicit symbolic real solutions of cubic equations. Such solutions can be efficiently substituted in real side conditions at practically low price of the linear and quadratic real quantifier elimination (Weispfenning, 1988, 1997) in REDLOG (Dolzmann and Sturm, 1997).

Example 13. We illustrate Theorem 9 using a simple example. Let

$$\begin{aligned}
&f(x) := x^3 - ax + 1\\
&\Gamma(x) := -1/2 \leq x \leq 1/2
\end{aligned}$$

where a is a parameter. Direct calculations, using the formula in Theorem 9, yield

$$\begin{aligned}
p_1 &= a_2^2 a_1^2 + 18\, a_2 a_1 a_0 - 4\, a_1^3 - 27\, a_0^2 - 4\, a_2^3 a_0 = 4\, a^3 - 27, \\
p_2 &= 9\, a_2 a_1 - 27\, a_0 - 2\, a_2^3 = -27, \\
s &= \sqrt[2]{-3\, p_1} = \sqrt[2]{81 - 12\, a^3}, \\
c_1 &= \sqrt[3]{(p_2 + 3\, s)/2} = \sqrt[3]{(-27 + 3\, \sqrt[2]{81 - 12\, a^3})/2}, \\
c_2 &= \sqrt[3]{(p_2 - 3\, s)/2} = \sqrt[3]{(-27 - 3\, \sqrt[2]{81 - 12\, a^3})/2}, \\
\sigma &= \mathrm{sign}(p_2) = -1, \\
u_1 &= (-a_2 + \omega^{(1-\sigma)} c_1 + \omega^{(2+\sigma)} c_2)/3 = (\omega^2 c_1 + \omega^1 c_2)/3, \\
u_2 &= (-a_2 + \omega^{(0-\sigma)} c_1 + \omega^{(0+\sigma)} c_2)/3 = (\omega^1 c_1 + \omega^2 c_2)/3, \\
u_3 &= (-a_2 + \omega^{(2-\sigma)} c_1 + \omega^{(1+\sigma)} c_2)/3 = (\omega^0 c_1 + \omega^0 c_2)/3,
\end{aligned}$$

and

$$\begin{aligned}
&\Gamma_j := (\exists\, x \in \mathbb{R})\, [x^3 - ax + 1 = 0 \wedge -1/2 \le x \le 1/2 \wedge \Phi_j(x)], \quad j = 1, 2, 3, \\
&\Phi_1(x) := [\, 3\, x^2 - a > 0 \wedge 6\, x > 0\,] \vee [\, 3\, x^2 - a = 0 \wedge 6\, x \ge 0\,], \\
&\Phi_2(x) := [\, 3\, x^2 - a \le 0\,] \qquad\qquad\quad \vee [\, 6\, x = 0\,], \\
&\Phi_3(x) := [\, 3\, x^2 - a > 0 \wedge 6\, x < 0\,] \vee [\, 3\, x^2 - a = 0 \wedge 6\, x \le 0\,].
\end{aligned}$$

Using the real quantifier elimination procedure QEPCAD (Collins and Hong, 1991; Brown and Hong, 2004) to eliminate the existential quantifiers in the above formula, we obtain the following quantifier-free formulas equivalent to Γ_j:

$$\begin{aligned}
&\Gamma_1 \iff false, \\
&\Gamma_2 \iff 4\, a - 9 \ge 0, \\
&\Gamma_3 \iff 4\, a + 7 \le 0.
\end{aligned}$$

Hence we finally obtain

$$[x^3 - ax + 1 = 0 \wedge -1/2 \le x \le 1/2] \iff [x = u_2 \wedge 4\, a - 9 \ge 0] \vee [x = u_3 \wedge 4\, a + 7 \le 0].$$

We can also use the real quantifier elimination function in REDLOG (Dolzmann and Sturm, 1997) to obtain the following quantifier-free formulas equivalent to Γ_j:

$$\begin{aligned}
&\Gamma_1 \iff false, \\
&\Gamma_2 \iff 4\, a^3 - 27 > 0 \wedge 4\, a - 9 \ge 0, \\
&\Gamma_3 \iff 4\, a^3 - 27 < 0 \wedge 4\, a + 7 \le 0.
\end{aligned}$$

Simplifying the above formulas, we get the same result as using QEPCAD. □

5. Proof of the correctness of the cubic formula with constraints

In this section, we prove Theorem 9 stated in the previous section. The proof will be divided into the following two lemmas. The proof of each lemma will be further divided into cases depending on the solution indexing in Figure 2.

Lemma 14. $u_1 = r_1$, $u_2 = r_2$, and $u_3 = r_3$.

Proof. We use the results and the same $q_i,\ q_i'$ from Lemma 7.

(1) $p_1 > 0$ and $p_2 > 0$. In this case, we have $c_1 = q_1,\ c_2 = q_1',\ \sigma = +1$. Substituting c_1 and c_2 into u_k in Theorem 9 and simplifying the resulting expressions using $\omega^3 = 1$ and $\omega^0 + \omega^1 + \omega^2 = 0$, we see that

$$\begin{aligned} u_1 &= \frac{3\,r_1 + (\omega^0+\omega^1+\omega^2)\,r_1 + (\omega^0+\omega^1+\omega^2)\,r_2}{3} = r_1, \\ u_2 &= \frac{3\,r_2 + (\omega^0+\omega^1+\omega^2)\,r_1 + (\omega^0+\omega^1+\omega^2)\,r_3}{3} = r_2, \\ u_3 &= \frac{3\,r_3 + (\omega^0+\omega^1+\omega^2)\,r_1 + (\omega^0+\omega^1+\omega^2)\,r_2}{3} = r_3. \end{aligned}$$

(2) $p_1 > 0$ and $p_2 = 0$. In this case $c_1 = q_2,\ c_2 = q_2',\ \sigma = 0$. Similar calculation yields $u_1 = r_1,\ u_2 = r_2,\ u_3 = r_3$.

(3) $p_1 > 0$ and $p_2 < 0$. In this case $c_1 = q_3,\ c_2 = q_3',\ \sigma = -1$. Similar calculation yields $u_1 = r_1,\ u_2 = r_2,\ u_3 = r_3$.

(4) $p_1 = 0$ and $p_2 > 0$. In this case $c_1 = q_1,\ c_2 = q_1',\ \sigma = +1$. Similar calculation yields $u_1 = r_1,\ u_2 = r_2,\ u_3 = r_3$.

(5) $p_1 = 0$ and $p_2 = 0$. In this case $c_1 = q_2,\ c_2 = q_2',\ \sigma = 0$. Similar calculation yields $u_1 = r_1,\ u_2 = r_2,\ u_3 = r_3$.

(6) $p_1 = 0$ and $p_2 < 0$. In this case $c_1 = q_3,\ c_2 = q_3',\ \sigma = -1$. Similar calculation yields $u_1 = r_1,\ u_2 = r_2,\ u_3 = r_3$.

(7) $p_1 < 0$ and $p_2 > 0$. In this case $c_1 = q_1,\ c_2 = q_1',\ \sigma = +1$. Similar calculation yields $u_1 = r_1,\ u_2 = r_2,\ u_3 = r_3$.

(8) $p_1 < 0$ and $p_2 = 0$. In this case $c_1 = q_2,\ c_2 = q_2',\ \sigma = 0$. Similar calculation yields $u_1 = r_1,\ u_2 = r_2,\ u_3 = r_3$.

(9) $p_1 < 0$ and $p_2 < 0$. In this case $c_1 = q_3,\ c_2 = q_3',\ \sigma = -1$. Similar calculation yields $u_1 = r_1,\ u_2 = r_2,\ u_3 = r_3$.

□

The indexing of solutions in Figure 2 also permits us to establish the following lemma using the idea underlying Thom's lemma (Basu et al., 2006, p. 50): each real r_k is uniquely determined by the signs of the derivatives of f at r_k.

Lemma 15. $\Gamma_j \iff r_j \in \mathbb{R} \wedge \Gamma(r_j)$.

Proof. Note that

$$\Gamma_j := (\exists\, z \in \mathbb{R})\,[f(z) = 0 \wedge \Gamma(z) \wedge \Phi_j(z)] \iff \bigvee\nolimits_{k=1}^{3} r_k \in \mathbb{R} \wedge \Gamma(r_k) \wedge \Phi_j(r_k).$$

We need to determine $\Phi_j(r_k)$. For this, observe that

$$
\begin{aligned}
f'(r_1) &= (r_1 - r_2)(r_1 - r_3), & f''(r_1) &= 2\,(r_1 - r_2) + 2\,(r_1 - r_3),\\
f'(r_2) &= (r_2 - r_1)(r_2 - r_3), & f''(r_2) &= 2\,(r_2 - r_1) + 2\,(r_2 - r_3),\\
f'(r_3) &= (r_1 - r_3)(r_2 - r_3), & f''(r_3) &= 2\,(r_3 - r_1) + 2\,(r_3 - r_2).
\end{aligned}
$$

For each configuration of the solutions, we can determine the signs of the derivatives of f at r_k, as in Table 1 (where the blanks are non-real). From the signs of the derivatives, it is easy to obtain the truth values of Φ_j as in Table 2 (where the blanks are false).

Table 1: Signs of derivatives of f

p_1	$+$	$+$	$+$	0	0	0	$-$	$-$	$-$
p_2	$+$	0	$-$	$+$	0	$-$	$+$	0	$-$
$f'(r_1)$	$+$	$+$	$+$	$+$	0	0	$+$	0	
$f''(r_1)$	$+$	$+$	$+$	$+$	0	$+$	$+$		
$f'(r_2)$	$-$	$-$	$-$	0	0	0		$+$	
$f''(r_2)$	$-$	0	$+$	$-$	0	$+$		0	
$f'(r_3)$	$+$	$+$	$+$	0	0	$+$		0	$+$
$f''(r_3)$	$-$	$-$	$-$	$-$	0	$-$			$-$

Table 2: Truth values of Φ_j

p_1	$+$	$+$	$+$	0	0	0	$-$	$-$	$-$
p_2	$+$	0	$-$	$+$	0	$-$	$+$	0	$-$
$\Phi_1(r_1)$	true	true	true	true	true	true	true		
$\Phi_1(r_2)$					true	true			
$\Phi_1(r_3)$					true				
$\Phi_2(r_1)$					true	true			
$\Phi_2(r_2)$	true	true	true	true	true	true		true	
$\Phi_2(r_3)$				true	true				
$\Phi_3(r_1)$					true				
$\Phi_3(r_2)$				true	true				
$\Phi_3(r_3)$	true	true	true	true	true	true			true

From Table 2, we see immediately that

$$\Gamma_j \iff \bigvee\nolimits_{k=1}^{3} r_k \in \mathbb{R} \wedge \Gamma(r_k) \wedge \Phi_j(r_k) \iff r_j \in \mathbb{R} \wedge \Gamma(r_j).$$

$\square$

Proof of Theorem 9. Let $x \in \mathbb{R}$. By Lemmas 14 and 15, we have

$$
\begin{aligned}
f(x) = 0 \wedge \Gamma(x) &\Longleftrightarrow (x = r_1 \vee x = r_2 \vee x = r_3) \wedge \Gamma(x) \\
&\Longleftrightarrow [x = r_1 \wedge r_1 \in \mathbb{R} \wedge \Gamma(r_1)] \vee \\
&\qquad [x = r_2 \wedge r_2 \in \mathbb{R} \wedge \Gamma(r_2)] \vee \\
&\qquad [x = r_3 \wedge r_3 \in \mathbb{R} \wedge \Gamma(r_3)] \\
&\Longleftrightarrow [x = u_1 \wedge \Gamma_1] \vee [x = u_2 \wedge \Gamma_2] \vee [x = u_3 \wedge \Gamma_3].
\end{aligned}
$$

The theorem is proved. □

6. Concluding remarks

We have presented the following:

- A real convention which provides correct interpretations of Lagrange's formula for all cubic polynomial equations with real coefficients;
- Real solution formulas for the general cubic equation $f = 0$ under equality and inequality constraints, in which the three real solutions are separated by using the signs of the first- and the second-order derivatives of f.

Yet the following questions still remain for future investigation.

- Whether there is a convention that yields correct solutions for all cubic polynomial equations with *complex* coefficients.
- Whether Theorem 9 and the result in Weispfenning (1994) can be combined to obtain a more efficient formulation. This insightful question was raised by an anonymous referee who also suggested that there should be a strong connection between the solutions u_i in the second part of the present paper and the symbolic solutions γ_i and real types of polynomials in Weispfenning (1994). We have investigated the issues and indeed there is a strong connection. However, we are not yet able to combine them into a better formulation due to various technical subtleties. We agree that it is worthwhile to pursue this as future work.
- How to generalize the solution formulas from the cubic to the quartic case. For this, one might need to carefully examine the theories underlying Sturm-Habicht sequences and discriminant systems (Gonzalez, et al., 1989; Yang et al., 1996; Yang and Xia, 1997; Liang and Zhang, 1999).
- How effective these formulas are for applications, in particular to dynamic geometric constraint solving (Hong et al., 2006).

Acknowledgements

The research of T. Zhao and D. Wang has been supported partially by the ANR-NSFC Project ANR-09-BLAN-0371-01/60911130369 (EXACTA) and the SKLSDE Open Fund BUAA-SKLSDE-09KF-01. The research of H. Hong has been supported partially by US NSF 0532140 (SCREMS). The authors would like to thank the anonymous referees for their insightful remarks and helpful suggestions (mentioned above).

References

Basu, S., Pollack, R., Roy, M. F., 2006. Algorithms in Real Algebraic Geometry. Springer-Verlag, New York.

Bortolotti, E. (Ed.), 1966. R Bombelli L'Algebra, Books I–V. Milan.

Brown, C. W., Hong, H., 2004. QEPCAD — Quantifier Elimination by Partial Cylindrical Algebraic Decomposition. http://www.cs.usna.edu/~qepcad/B/QEPCAD.html.

Cardano, G., 1993. Ars Magna or The Rules of Algebra. Dover Publications, New York.

Collins, G. E., Hong, H., 1991. Partial Cylindrical Algebraic Decomposition for Quantifier Elimination. Journal of Symbolic Computation 12, 99–328.

Dolzmann, A., Sturm, T., 1997. Smplification of Quantifier-free Formulas over Ordered Fields. Journal of Symbolic Computation 24 (2), 209–231.

Dolzmann, A., Sturm, T., 1997. REDLOG, Computer Algebra Meets Computer Logic. ACM SIGSAM Bulletin 31 (2), 2–9.

Gonzalez, L., Lombardi, H., Recio, T., Roy, M. F., 1989. Sturm-Habicht Sequence, in: Gonnet G. H. (Ed.), Proceedings of the ACM-SIGSAM 1989 International Symposium on Symbolic and Algebraic Computation. ACM Press, New York, pp. 136–146.

Guilbeau, L., 1930. The History of the Solution of the Cubic Equation. Mathematics News Letter 5 (4), 8–12.

Lagrange, J. L., 1770. Réflexions sur la résolution algébrique des équations, in: Nouveaux Mémoires de l'Académie royale des Sciences et Belles-Lettres de Berlin, Oeurves complètes, tome 3, pp. 205–421.

Liang, S., Zhang, J., 1999. A Complete Discrimination System for Polynomials with Complex Coefficients and Its Automatic Generation. Science in China (Ser. E) 42 (2), 113–128.

Hong, H., Li, L., Liang, T., Wang, D., 2006. Solving Dynamic Geometric Constraints Involving Inequalities, in: Calmet J., Ida T., Wang D. (Eds.), Artificial Intelligence and Symbolic Computation. Springer-Verlag, Berlin Heidelberg, LNAI 4120, pp. 181–195.

O'Connor, J. J., Robertson, E. F., 2010. Rafael Bombelli, in: MacTutor History of Mathematics Archive. Unversity of St. Andrew, Soctland.

Smith, D. E., 2003. History of Mathematics (Vol. 2). Dover Publications, New York.

Weispfenning, V., 1988. The Complexicity of Linear Problems in Fields. Journal of Symbolic Computation 5, 3–27.

Weispfenning, V., 1994. Quantifier Elimination for Real Algebra — the Cubic Case, in: Proceedings of the International Symposium on Symbolic and Algebraic Computation. ACM Press, New York, pp. 258–263.

Weispfenning, V., 1997. Quantifier Elimination for Real Algebra — the Quadratic Case and Beyond. Applicable Algebra in Engineering, Communication and Computing 8, 3–27.

Wang, D., 2004. GEOTHER 1.1: Handling and Proving Geometric Theorems Automatically, in: Winkler F. (Ed.), Automated Deduction in Geometry. Springer-Verlag, Berlin Heidelberg, LNAI 2930, pp. 194–215.

Yang, L., Hou, X., Zeng, Z., 1996. A Complete Discrimination System for Polynomials. Science in China (Ser. E) 39 (6), 628–646.

Yang, L., Xia, B., 1997. Explicit Criterion to Determine the Number of Positive Roots of a Polynomial. MM Research Preprints 15, 134–145.