

Cubic polynomials with real or complex coefficients: The full picture

Nicholas S. Bardell

RMIT University

nick.bardell@rmit.edu.au

Introduction

The cubic polynomial with real coefficients $y = ax^3 + bx^2 + cx + d$ in which $a \neq 0$, has a rich and interesting history primarily associated with the endeavours of great mathematicians like del Ferro, Tartaglia, Cardano or Vieta who sought a solution for the roots (Katz, 1998; see Chapter 12.3: The Solution of the Cubic Equation). Suffice it to say that since the times of renaissance mathematics in Italy various techniques have been developed which yield the three roots of a general cubic equation. A ‘cubic formula’ does exist—much like the one for finding the two roots of a quadratic equation—but in the case of the cubic equation the formula is not easily memorised and the solution steps can get quite involved (Abramowitz & Stegun, 1970; see Chapter 3: Elementary Analytical Methods, 3.8.2 Solution of Cubic Equations). Hence it is not surprising that with the advent of the digital computer, numerical root-finding algorithms such as those attributed to Newton-Raphson, Halley, and Householder have become the solution of choice (Weisstein, n.d.).

Students studying a mathematics specialism as a part of the Victorian VCE (Mathematical Methods, 2010), HSC in NSW (Mathematics Extension in NSW, 1997) or Queensland QCE (Mathematics C, 2009) are expected to be able to recognise, plot, factorise, and even solve cubic polynomials as described by ACARA (ACARA, n.d., Unit 1, Topic 1: Functions and Graphs). The latter requirement even goes so far as to encourage students to “solve cubic equations using technology”. Since the application of such technology, e.g., MathCAD (2007), is fundamental to the contents of this paper, it is hoped that the work presented here will prove interesting, relevant, and stimulating to students and teachers alike.

It is not the intention here to present a new method for finding the roots of a cubic polynomial; rather, the aim is to allow all the roots to be visualised, their arrangement better understood, and an explanation provided for the complex conjugate root pairing. This understanding becomes of paramount importance when the coefficients are generally complex and any combination

of roots can result; an explanation for this latter case appears to be completely absent from the literature.

Cubic polynomials with real coefficients

To begin, recall a quadratic function $y = ax^2 + bx + c$ can *always* be cast in vertex form $y = a(x - h)^2 + k$ for some constants h and k which means the graph of $y = a(x - h)^2 + k$ can be obtained from the standard parabola $y = x^2$ by reflections, dilations and translations. However, this is not so for cubic polynomials since not every cubic graph can be obtained from the standard cubic $y = x^3$ by a combination of reflections, dilations and translations. This failure/inability to express a cubic polynomial in vertex form tends to complicate matters considerably and will be shown to have quite an impact on the following investigation.

Consider the general cubic equation

$$y = ax^3 + bx^2 + cx + d \quad (1)$$

in which the coefficients a , b , c and d are real and $a \neq 0$. For completeness, a brief review of cubic polynomials is now given; these salient facts can be found in any good textbook and are presented here without proof.

Figure 1 illustrates the three most general forms a cubic function may assume. The increased complexity compared with a quadratic function should immediately be evident.

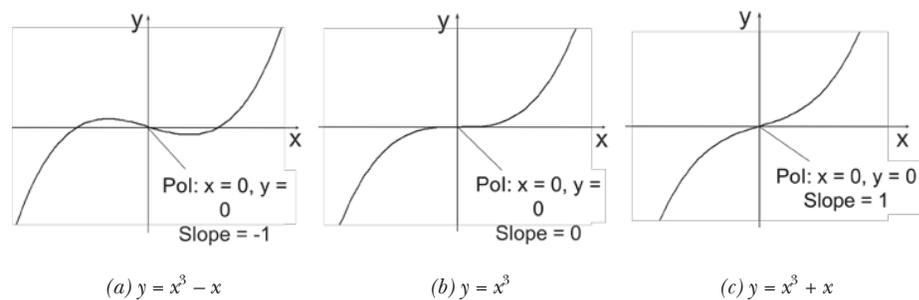


Figure 1. Illustration of the variation that is possible in cubic curves (PoI = Point of Inflection).

In order to produce an accurate graph of a given cubic function, it is necessary to identify four critical points of interest:

1. the location of the relative extremes;
2. the location at which the sign of the curvature (concavity) changes (i.e., the point of inflection (PoI));
3. the location of the point(s) where the graph crosses the x -axis (i.e., the roots); and
4. the location of the point where the graph crosses the y -axis (i.e., the y -intercept).

The location of the turning points

The turning points are found by differentiating Equation (1) and setting the resulting equation for the slope equal to zero, that is,

$$\frac{dy}{dx} = 3ax^2 + 2bx + c = 0 \quad (2)$$

The location of the turning points is found by using the quadratic formula to solve Equation (2):

$$x = \frac{-b \pm \sqrt{b^2 - 3ac}}{3a} \quad (3)$$

1. If $b^2 - 3ac > 0$, two distinct turning points result—characterised by a local maximum and a local minimum—giving the classic S-shaped cubic curve. For an example, see Figure 1(a). Note that, unlike quadratic curves, the turning points of a cubic are not symmetrically located between x -axis intercepts.
2. If $b^2 - 3ac = 0$, there are no distinct turning points; they have effectively coalesced into a single stationary point of inflection. For an example, see Figure 1(b).
3. If $b^2 - 3ac < 0$, there are no distinct turning points and the function is monotonic. For an example, see Figure 1(c).

Further clarification about the nature of the turning points can be found by checking the value of the slope either side of the turning point. Alternatively, the second derivative of Equation (1) may be used:

$$\text{If } \frac{d^2y}{dx^2} = 6ax + 2b \begin{cases} < 0 \text{ the turning point is a maximum} \\ > 0 \text{ the turning point is a minimum} \end{cases}$$

The location of the point of inflection

Every cubic polynomial has one point of inflection (PoI), that is, a point on the curve at which the sign of the curvature (i.e., the concavity) changes. The PoI is found by setting the second derivative of Equation (1) equal to zero thus:

$$\frac{d^2y}{dx^2} = 6ax + 2b = 0, \text{ from which } x = -\frac{b}{3a} \quad (4)$$

If, in addition to

$$\frac{d^2y}{dx^2} = 0$$

the third derivative of Equation (1) is non-zero, the existence of a point of inflection is confirmed. This is always the case for a cubic polynomial because

$$\frac{d^3y}{dx^3} = 6a$$

and $6a \neq 0$ since by definition $a \neq 0$. Hence

$$x = -\frac{b}{3a}$$

always defines the location of the PoI on a cubic polynomial.

It is further noted that:

- A PoI may be a stationary point, like the one shown in Figure 1(b), but it is not a local maximum or a local minimum.
- A PoI is not necessarily parallel to the x -axis. An example of this is shown in Figure 1(c).
- If distinct turning points are present, like the ones shown in Figure 1(a), the PoI also corresponds to the mid-point between them. This is evident from Equation (3), which confirms the turning points are located equal distances either side of $x = -b/3a$, which is the location of the PoI.

The location of the roots

A cubic equation has three roots which are found by setting $y = 0$ in Equation (1) and solving for x . Since the leading coefficient is not zero, then

$$\lim_{x \rightarrow -\infty} ax^3 + bx^2 + cx + d = \begin{cases} -\infty & \text{if } a > 0 \\ +\infty & \text{if } a < 0 \end{cases}$$

$$\lim_{x \rightarrow +\infty} ax^3 + bx^2 + cx + d = \begin{cases} +\infty & \text{if } a > 0 \\ -\infty & \text{if } a < 0 \end{cases}$$

Since a polynomial is continuous, then by the Intermediate Value Theorem, if it takes a positive value and a negative value, it must take every value in between, in particular zero. Hence the graph of a cubic must cross the x -axis at least once, and thus one of the roots is always real. The arrangement of the roots can occur in any of the following combinations:

- one real root and a pair of complex conjugate roots.
- three real roots:
 - all of which may be coincident;
 - one of which may be distinct and the remaining pair coincident;
 - all of which may be distinct.

The location of the y-intercept

The y -intercept is found by setting $x = 0$ in Equation (1). This immediately gives $y_{\text{intercept}} = d$.

Cubic polynomials in the Cartesian plane

To illustrate many of the features discussed above for cubic polynomials with real coefficients plotted in the Cartesian plane, consider the curve $y = x^3 + 3x^2 + x - 5$ shown in Figure 2 over the range $-4 \leq x \leq 2$. There is one real root apparent at $x = 1$ but no indication whatsoever of the whereabouts of the complex conjugate roots. In this particular instance it is easy enough to divide the expression for y by the known factor $(x - 1)$ to yield $y = (x - 1)(x^2 + 4x + 5)$ from which the remaining two roots are found to be $-2 \pm i$. There are two distinct turning points present which, from Equations (3) and (1), are found to occur at $(x, y) = (0.184, -5.089)$ MIN and $(-1.816, -2.911)$ MAX respectively. Using Equations (4) and (1), the PoI is seen to occur at

$(x, y) = (-1, -4)$, midway between the turning points. The y -intercept is found at $y = -5$.

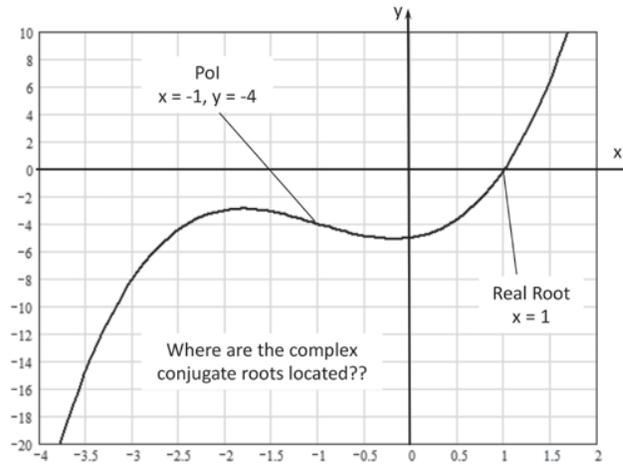


Figure 2. Plot of the curve $y = x^3 + 3x^2 + x - 5$ over the range $-4 \leq x \leq 2$.

The complex conjugate roots do not correspond to the locations of either turning point or the PoI shown in Figure 1 and have often been misplaced by well-meaning but ingenuous authors like Stroud (1986, see Programme 2, Theory of Equations, where he [incorrectly] places the complex conjugate roots in the Cartesian plane at one of the turning points of a cubic equation). Clearly, a two-dimensional representation is incapable of showing the full picture; the main motivation of this paper is showing how this limitation can be overcome. A full answer to questions like: “Where exactly are the roots located, then?” and “Why can’t I see them?” is given here along with some further discussion about what happens in the case of complex coefficients.

Visualising the bigger picture

Since two of the roots resulting from Equation (1) can be complex, this implies that the values assigned to x in Equation (1) do not have to be limited to the set of real numbers but rather could—and possibly should—include complex values. Whilst this is implicitly obvious, since no restriction was ever placed on x or y at the outset, it is rarely presented as an explicit proposition. Perhaps one reason why this is seldom considered is simply one of expedience—constructing a plot using generally complex values of x is not for the faint-hearted! Nevertheless, if this line of reasoning is pursued, and x is allowed to take a generally complex value $G + iH$, in which i is the complex number $\sqrt{-1}$, then the general cubic polynomial becomes:

$$y = a(G + iH)^3 + b(G + iH)^2 + c(G + iH) + d \quad (5a)$$

Equation (5a) can be expanded and simplified to:

$$y = aG^3 - 3aG^2H + bG^2 - bH^2 + cG + d + i(3aG^2H - aH^3 + 2bGH + cH) \quad (5b)$$

Since y must also be generally complex, of the form $A + iB$, then it is possible to equate the real (Re) and imaginary (Im) parts of both sides of Equation (5b).

$$\text{Equating Re parts: } A = aG^3 - 3aG^2H + bG^2 - bH^2 + cG + d \quad (6)$$

$$\text{Equating Im parts: } B = 3aG^2H - aH^3 + 2bGH + cH \quad (7)$$

Equation (6) represents a surface in three-dimensional space, in which the original x - and y -axes are now labelled $\text{Re}(x) (\equiv G)$ and $\text{Re}(y) (\equiv A)$, and the axis in the third dimension is $\text{Im}(x) (\equiv H)$. Equation (7) also represents a surface in three-dimensional space, in which the original x - and y -axes are now labelled $\text{Re}(x) (\equiv G)$ and $\text{Im}(y) (\equiv B)$, and the axis in the third dimension is $\text{Im}(x) (\equiv H)$. Hence the horizontal plane contains the $\text{Re}(x) (\equiv G)$ and $\text{Im}(x) (\equiv H)$ axes, thus forming the Argand plane, whilst the vertical plane given by GA or GB represents the Cartesian plane for each surface. By calculating both individual surfaces defined by Equation (6) and Equation (7) for a range of complex x values, i.e., a range of corresponding G and H values, the true and most general form of Equation (1) can be visualised.

All the three-dimensional plots presented in this paper were constructed using Mathcad (2007), which is one of many VCE/QCE/HSC-approved computer algebra systems available to schools, and fully commensurate with the following ACARA (n.d., Rationale) stated aims “to develop students’ understanding and ability to solve applied problems using concepts and techniques drawn from geometry and complex numbers” and “to develop students’ capacity to choose and use technology appropriately”.

Cubic polynomials with real coefficients: Discussion example

To illustrate the salient features of the surface plots that result from Equation (6) and Equation (7) the curve presented in Figure 2 for which $y = x^3 + 3x^2 + x - 5$ is revisited. Since in this example the coefficients of the cubic equation are identified as: $a = 1$, $b = 3$, $c = 1$ and $d = -5$, then Equation (6) and Equation (7) become:

$$A = G^3 - 3G^2H + 3G^2 - 3H^2 + G - 5 \quad (8)$$

$$B = 3G^2H - H^3 + 6GH + H \quad (9)$$

By supplying a suitable range of values for H and G , such as $-4 \leq G \leq 2$ and $-3 \leq H \leq 3$ for this particular example, the three-dimensional surfaces A and B can be calculated and plotted as shown in Figures 3(a)–3(d) and 4(a)–4(d) respectively. A horizontal plane has also been inserted at zero altitude in both surface plots to aid visualisation.

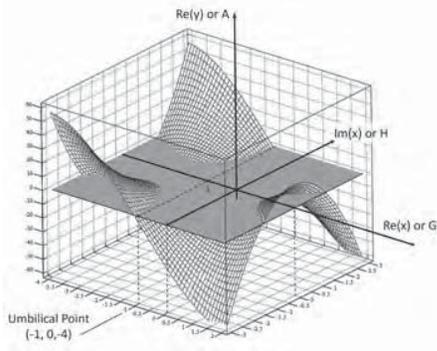


Figure 3(a). Surface A: The real part of the complex representation of $y = x^3 + 3x^2 + x - 5$.

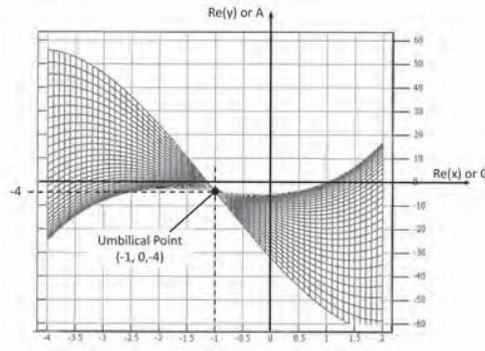


Figure 3(b). Side elevation of Figure 3(a) looking on the GA plane.

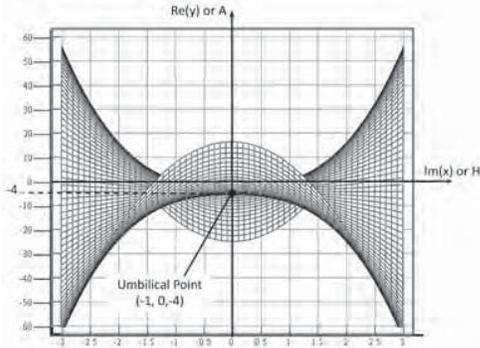


Figure 3(c). Side elevation of Figure 3(a) looking on the HA plane.

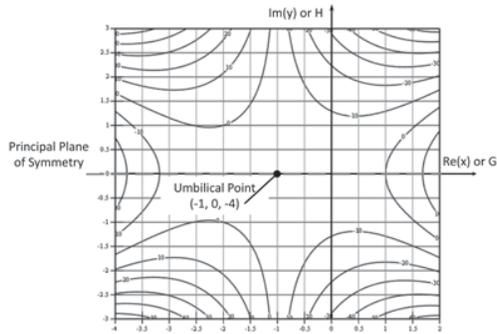


Figure 3(d). Plan view of Figure 3(a) looking on the GH (Argand) plane.

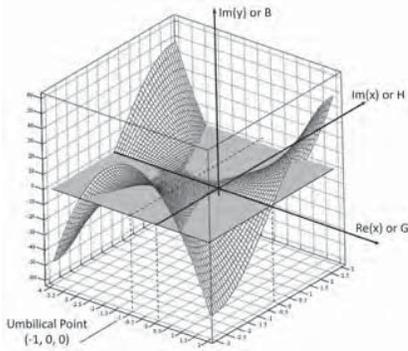


Figure 4(a). Surface B: The imaginary part of the complex representation of $y = x^3 + 3x^2 + x - 5$.

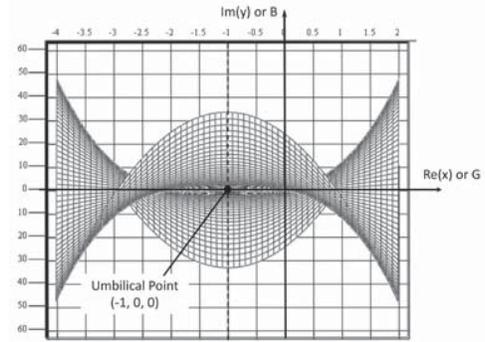


Figure 4(b). Side elevation of Figure 4(a) looking on the GB plane.

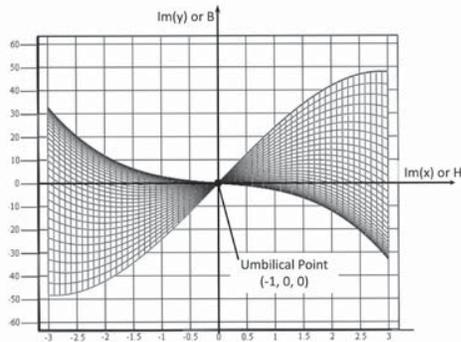


Figure 4(c). Side elevation of Figure 4(a) looking on the HB plane.

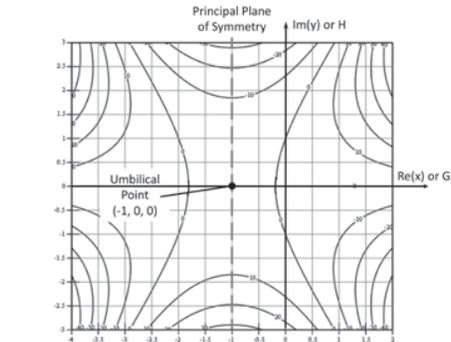


Figure 4(d). Plan view of Figure 4(a) looking on the GH (Argand) plane.

Surfaces A and B each belong to a class of surfaces termed a *monkey saddle* since a saddle for a monkey requires three depressions: two for the legs, and one for the tail. For each surface, the principal plane of symmetry is identified as the symmetry plane that aligns with, or runs parallel to, the G - or H -axis respectively; see Figures 3(d) and 4(d).

At this point the reader should observe that there are different mathematical categories of three-dimensional surfaces, each with its own set of distinctive properties and characteristics.

- A surface of order one (e.g., $ax + by + cz + d = 0$) is called a plane.
- A surface of order two is called a quadratic or quadric surface and consists of surfaces such as ellipsoids, paraboloids and hyperboloids, which include the classic MINIMAX saddle surfaces.
- A surface of order three, otherwise known as a cubic surface, is the type of surface studied in this paper, since the highest degree of any term, or combination of terms, shown in Equations (6) and (7) is three. These singular cubic surfaces were first investigated and classified by Schläfli (1863) and Cayley (1869).

As mentioned previously, a cubic polynomial cannot, in general, be cast in vertex form. This has a significant bearing on the current problem, since the technique presented by Bardell (2012, 2014a) that enabled a quadratic surface to be rendered in its canonical form using various translations and rotations no longer works for a cubic surface. Although it is thus not possible to cast the cubic surfaces defined by Equations (6) and (7) in a simplified form about some new origin, these surfaces may nonetheless still be plotted without loss of generality. Since it is beyond the scope of this paper to delve into the theory of cubic surfaces, the reader is asked to accept that there is one important reference point that conveniently serves the purpose of a local origin, otherwise known as an *umbilical* point. This point is indicated in Figures 3–5 and is a localised point of inflection that can exist on a three dimensional surface; what makes this point unique, and hence a convenient reference point, is that it has zero Gaussian curvature in any direction (Harvey Mudd, n.d.). In just the same way that there is only one PoI on a cubic curve, there is also only one unique umbilical point on a cubic surface.

Readers seeking a more advanced understanding of surfaces are referred to Lipschulz (1969) and Falconer (1986) who give a fuller treatment of Gauss's work on surface curvature and introduce Riemann's extension to higher-dimensional spaces called manifolds.

With reference to Figures 3(d) and 4(d), when viewed from above, Surfaces A and B both share a common umbilical point located in the Argand plane at $G = -1$, $H = 0$. However, as can be seen from Figure 3(b) and 4(b), the vertical location of the umbilical point is unique to each surface and differs between them. The umbilical point for Surface B is always fixed at $B = 0$ since this is the outcome of substituting $H = 0$ into Equation (7). For Surface A the vertical location of the umbilical point is found by substituting $G = -1$ and $H = 0$ into

Equation (8) to give a value of $A = -4$. Clearly, Surface A (and its umbilical point) is free to translate vertically (up or down) along a normal to the GH plane passing through the umbilical point. Similar behaviour is observed for a general cubic polynomial when plotted in the Cartesian plane—for a given set of coefficients, a , b , c define the shape of the curve whilst the constant term d dictates its vertical positioning.

A section taken through Surface A at $H = 0$ reveals the trace of the original cubic polynomial in the GA (or Cartesian) plane. This follows from setting $H = 0$ in Equation (6), giving the curve $A = aG^3 + bG^2 + cG + d$. Figure 5 shows such a section for the current example (see Equation (8)) which clearly reveals the trace of the cubic polynomial $y = x^3 + 3x^2 + x - 5$ originally presented in Figure 2. This vividly illustrates how the original Cartesian x - y plane is merely a ‘slice’ of a much bigger picture. Figure 5 also shows the coincidence of the umbilical point on Surface A with the PoI on the original curve; these points represent the same notion of zero curvature in three- and two-dimensions respectively.

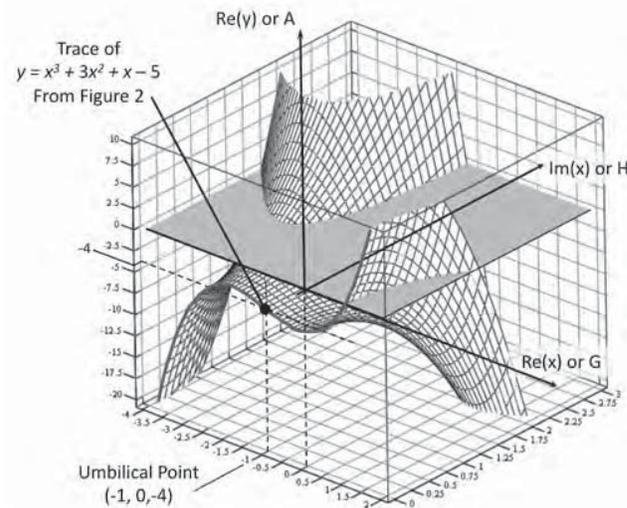


Figure 5. Section through the surface $A = G^3 - 3G^2H + 3G^2 - 3H^2 + G - 5$ taken at $H = 0$, showing the trace of the cubic polynomial $y = x^3 + 3x^2 + x - 5$ originally presented in Figure 2.

It is further observed that the surface B has no influence at all on the GA (or Cartesian) plane. This follows from setting $H = 0$ in Equation (7), giving the curve $B = 0$. This proves that the surface B always assumes a straight ‘nodal’ line coincident with the G -axis as it passes through $H = 0$; see Figure 4(a). This explains why the single curve shown in Figure 2 provides a full and complete description of $y = x^3 + 3x^2 + x - 5$. This observation holds for all cubic polynomials with real coefficients—however, it does not apply if the coefficients are generally complex for reasons which will be described later in the section entitled *Cubic polynomials with complex coefficients: General observations*.

Cubic polynomials with real coefficients: General observations

(presented without proof)

- The cubic form of surface A is identified as a general monkey saddle on account that a monkey sitting astride the G -axis of surface A has a place for its two legs and tail.

- Surface A has an umbilical point located at

$$G_U = -\frac{b}{3a}, H_U = 0, A_U = \frac{2b^3 - 9abc + 27a^2d}{27a^2}$$

- The principal vertical plane of symmetry of surface A always coincides with the GA plane at $H = 0$.
- Surface A is free to translate vertically along a normal to the GH (or Argand) plane through its umbilical point on account of the constant term d in Equation (6).

- The cubic form of surface B may also be identified as a general monkey saddle because a monkey sitting astride an axis through the umbilical point parallel to the H -axis of surface B has a place for its two legs and tail.

- Surface B has an umbilical point located at

$$G_U = -\frac{b}{3a}, H_U = 0, B_U = 0$$

- The principal vertical plane of symmetry of surface B is parallel to the HB plane passing through the point $G = -\frac{b}{3a}$.
- The umbilical point of surface B is fixed in the vertical sense at $B = 0$ which in turn fixes surface B . This is a direct consequence of there being no constant term in Equation (7).
- The principal vertical planes of symmetry of surfaces A and B are always at right angles to one another and, when viewed directly from above, always intersect through the umbilical point.

Cubic polynomials with real coefficients: Solution for the roots

Clearly, from the definition of a root to occur when $y = 0$, this implies that both $\text{Re}(y)$ and $\text{Im}(y)$ must simultaneously be zero. In other words, the location and nature of the roots will be defined where the two surfaces for $\text{Re}(y)$ ($\equiv A$) and $\text{Im}(y)$ ($\equiv B$) have common intersections with a horizontal plane positioned at zero altitude. As shown in Figure 6, there are three such common intersection points indicating the location of the three roots. For the specific example considered here, Figures 6 and 7 show one real root and a pair of complex conjugate roots located respectively at (G, H) or $(\text{Re}(x), \text{Im}(x)) = (1, 0)$, $(-2, 1)$ and $(-2, -1)$ which accords fully with the solution for $y = x^3 + 3x^2 + x - 5$, given by $x = 1, -2 \pm i$, described at the end of the section entitled *Cubic polynomials in the Cartesian plane*.

The manner in which the principal planes of symmetry of surfaces A and B intersect each other at right angles through the umbilical point helps explain why the complex conjugate roots have to occur in pairs equidistant from

the G -axis; they can now be visualised occurring behind and in front of the original Cartesian x - y plane.

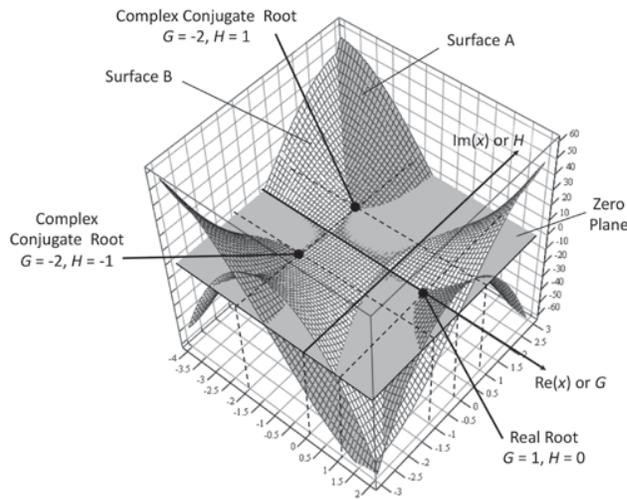


Figure 6. Visualization of the roots of $y = x^3 + 3x^2 + x - 5$ arising from the intersection of the surfaces $A = B = 0$ from the complex generalisation of the problem.

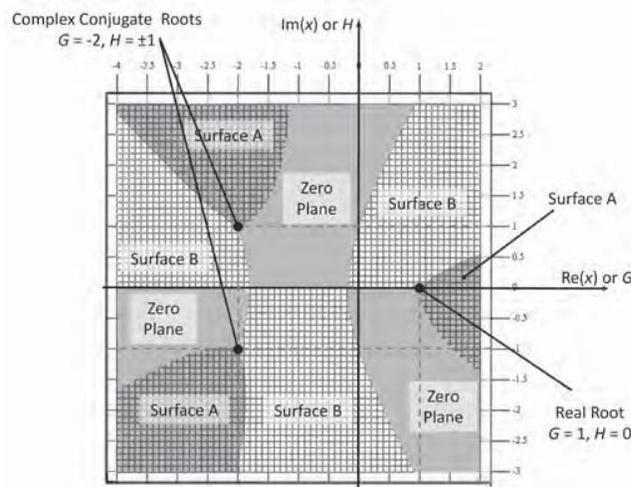


Figure 7. Plan view of Figure 6 showing the location of the roots of $y = x^3 + 3x^2 + x - 5$.

Cubic polynomials with real coefficients: Further examples

Figures 8, 9 and 10 illustrate how different cubic polynomials produce different root arrangements resulting from the intersections of their representative surfaces A and B . Figure 8(a) shows another instance where the complex conjugate roots are not evident in the Cartesian plane. The shape of this cubic differs from that shown in Figure 1 in that there are no turning points, the PoI is positively inclined to the x -axis, and the function is monotonically increasing with increasing x . Figures 8(b) and 8(c) show the full picture which immediately sheds light on the arrangement of the roots. Figures 9(b), 9(c), 10(b) and 10(c) show what really happens when only real roots result, with the latter case illustrating the phenomenon of a repeated real root. The

sharp intersection and the ‘petal-like’ arrangement of the surfaces on the right hand side of Figure 10(c) clearly shows how the coincident pair of roots is formed.

With reference to Figures 7, 8(c), 9(c) and 10(c), it is apparent that the roots

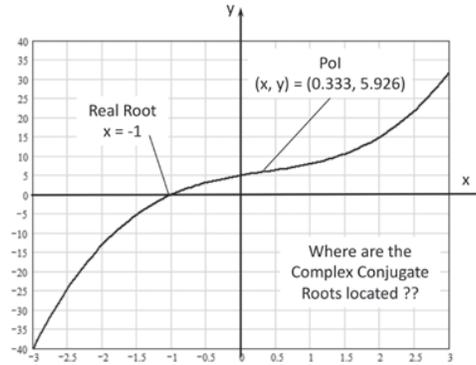


Figure 8(a). Plot of $y = x^3 - x^2 + 3x + 5$.

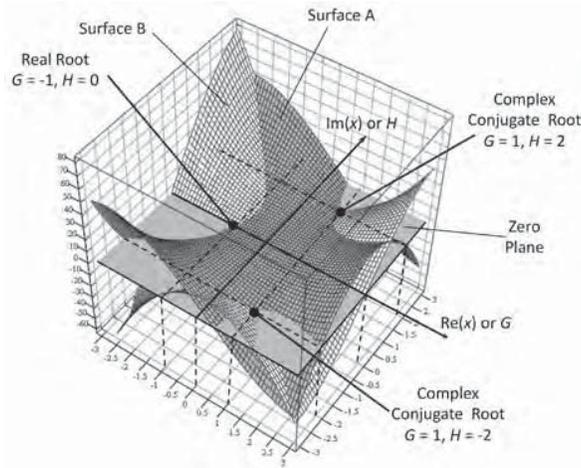


Figure 8(b). Visualisation of the roots of $y = x^3 - x^2 + 3x + 5$ arising from the intersection of the surfaces $A = B = 0$ from the complex generalisation of the problem.

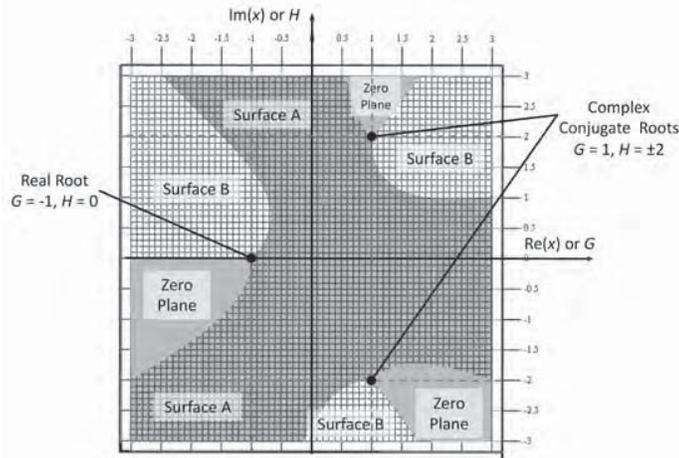


Figure 8(c). Plan view of Figure 8(b) showing the location of the roots of $y = x^3 - x^2 + 3x + 5$.

of a cubic with real coefficients always lie at the vertices of an isosceles triangle in the GH (or Argand) plane. The triangle is isosceles because one root is always on the horizontal (real) axis and the other two roots, being complex conjugates, appear symmetrically behind and in front of the real axis. When no complex conjugate roots are present, the isosceles triangle ‘collapses’ into a straight line; see Figures 9(c) and 10(c). Only in the special case when the cubic polynomial can be expressed in vertex form as $y = a(x - h)^3 + k$ does the isosceles triangle become equilateral and the roots then lie at equal 120° intervals around the circumscribed circle. Bardell (2014b) has provided some illustrations of this in the context of de Moivre’s Theorem and the n -th roots of unity.

Figure 8 shows the curve $y = x^3 - x^2 + 3x + 5$ and its generalisation in the complex domain. The PoI for the curve is located at $(x, y) = (0.333, 5.926)$; the umbilical for surface A is found at $(G, H, A) = (0.333, 0, 5.926)$. There are no turning points and the curve increases monotonically with increasing x . There is one real root at $x = -1$, and a complex conjugate pair at $x = 1 \pm 2i$.

Figure 9 shows the curve $y = x^3 - 4x^2 + x + 6$ and its generalisation in the complex domain. The PoI for the curve is located at $(x, y) = (1.333, 2.593)$; the umbilical for surface A is found at $(G, H, A) = (1.333, 0, 2.593)$. The turning points are found at $(x, y) = (0.131, 6.065)$ MAX and $(2.535, -0.879)$ MIN. There are three distinct real roots at $x = -1, 2, 3$.

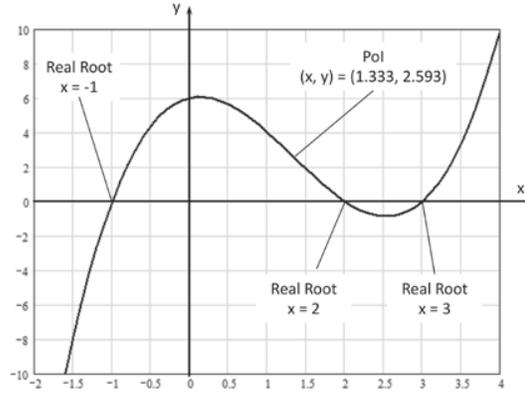


Figure 9(a). Plot of $y = x^3 - 4x^2 + x + 6$.

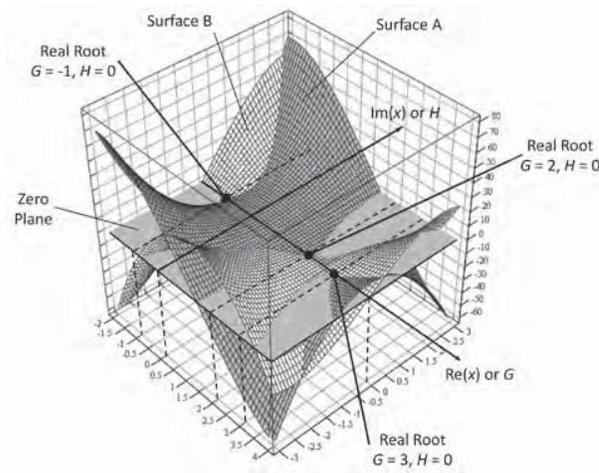


Figure 9(b). Visualization of the roots of $y = x^3 - 4x^2 + x + 6$ arising from the intersection of the surfaces $A = B = 0$ from the complex generalisation of the problem.

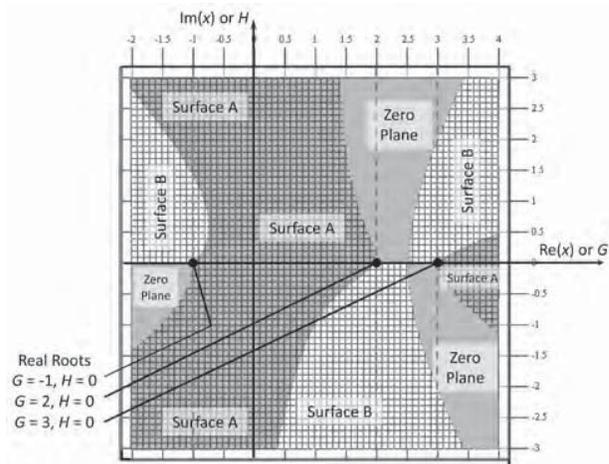


Figure 9(c). Plan view of Figure 9(b) showing the location of the roots of $y = x^3 - 4x^2 + x + 6$.

Figure 10 shows the curve $y = x^3 - 3x^2 + 4$ and its generalisation in the complex domain. The PoI for the curve is located at $(x, y) = (1, 2)$; the umbilical for surface A is found at $(G, H, A) = (1, 0, 2)$. The turning points are found at $(x, y) = (0, 4)$ MAX and $(2, 0)$ MIN. There are three real roots: one at $x = -1$, and a repeated pair at $x = 2$.

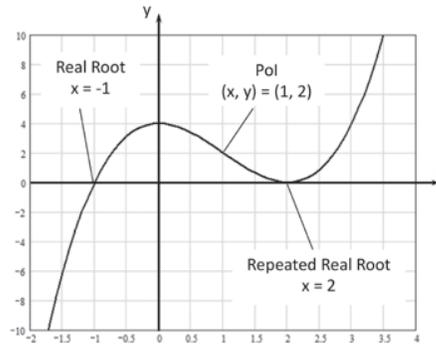


Figure 10(a). Plot of $y = x^3 - 3x^2 + 4$.

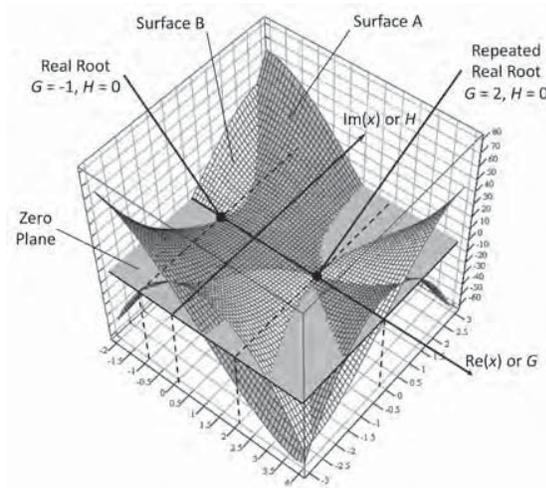


Figure 10(b). Visualization of the roots of $y = x^3 - 3x^2 + 4$ arising from the intersection of the surfaces $A = B = 0$ from the complex generalisation of the problem.

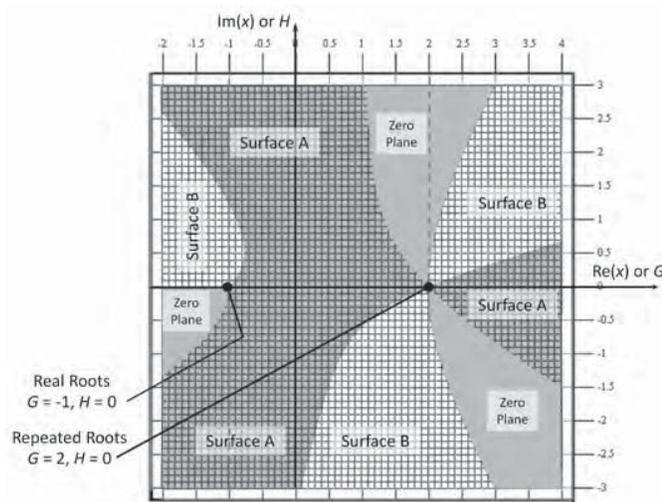


Figure 10(c). Plan view of Figure 10(b) showing the location of the roots of $y = x^3 - 3x^2 + 4$.

Cubic polynomials with generally complex coefficients

By permitting the coefficients to be generally complex, Equation (1) now assumes the form:

$$y = (a_R + ia_I)x^3 + (b_R + ib_I)x^2 + (c_R + ic_I)x + (d_R + id_I) \quad (10)$$

in which a_R, b_R, c_R and d_R are the real parts and a_I, b_I, c_I and d_I are the imaginary parts of the participating coefficients and either $a_R \neq 0$ or $a_I \neq 0$.

It is fair to say that it is hard to find anything substantive written about this topic. There are a few highly mathematical papers which consider cubic polynomials with complex coefficients at a conceptual level, e.g., Oswald (2009), Derdzinski (2014), but no worked examples or any commentary are provided. Whilst the present limitations of space and scope do not permit an in-depth discussion of this case, it is hoped the results, observations, and visualisations presented here will contribute something of merit to this topic.

As far as determining the roots is concerned, every method that is available for cubic polynomials with real coefficients is still applicable when the coefficients are complex, notwithstanding the increased difficulties associated with the calculations. Any combination of roots may now occur—the most common case is when all three roots are complex. Some of the numerical root-finding algorithms may require the user to input a guess value of the starting point for a given root, and this could easily be furnished by the visualisation techniques presented here.

A reasonable starting point would focus on plotting the curve represented by Equation (10). This is more difficult than it looks, and a simple x - y plot can only be found by separating the real and imaginary components of y and presenting them both on the same set of axes. However, such a plot is hard to interpret and in general does not reveal any information about the location or the nature of the roots. See Bardell (2014a) for a detailed discussion about the analogous case for a quadratic polynomial with complex coefficients—the same reasoning applies to the cubic case. In essence, because the slice of the surface represented by the Cartesian x - y plane will, in general, no longer contain any of the roots, a simple two-dimensional representation is destined to failure as far as their location is concerned.

As far as visualisation is concerned, the three-dimensional generalisation of a cubic equation with complex coefficients requires minimal extra effort and simply follows the same procedure as that described earlier for the real coefficient case. Now, by allowing x to take on the generally complex identity $G + iH$, Equation (10) becomes:

$$y = (a_R + ia_I)(G + iH)^3 + (b_R + ib_I)(G + iH)^2 + (c_R + ic_I)(G + iH) + (d_R + id_I) \quad (11a)$$

Equation (11a) can be expanded and simplified to:

$$y = a_R G^3 - 3a_R G H^2 - 3a_I G^2 H + a_I H^3 + b_R G^2 - b_R H^2 - 2b_I G H + c_R G - c_I H + d_R + i(3a_R G^2 H - a_R H^3 + a_I G^3 - 3a_I G H^2 + 2b_R G H + b_I G^2 - b_I H^2 + c_R H + c_I G + d_I) \quad (11b)$$

Since y must also be generally complex, of the form $A + iB$, then it is possible to equate the real (Re) and imaginary (Im) parts of both sides of Equation (11b).

Equating Re parts:

$$A = a_R G^3 - 3a_R G H^2 - 3a_I G^2 H + a_I H^3 + b_R G^2 - b_R H^2 - 2b_I G H + c_R G - c_I H + d_R \quad (12)$$

Equating Im parts:

$$B = 3a_R G^2 H - a_R H^3 + a_I G^3 - 3a_I G H^2 + 2b_R G H + b_I G^2 - b_I H^2 + c_R H + c_I G + d_I \quad (13)$$

From Equations (12) and (13) it is evident surfaces A and B are still cubic surfaces, albeit with a few more terms than the analogous case with real coefficients, and retain the shape of generalised monkey saddles. The roots are still found from the intersection of these surfaces with a plane at zero altitude. The major features introduced by complex coefficients are summarised here:

Cubic polynomials with generally complex coefficients:

General observations (presented without proof)

1. Surfaces A and B still share a common umbilical point when viewed from above in the Argand plane, but this is now located at:

$$G_U = -\frac{1}{3} \frac{(a_I b_I + a_R b_R)}{(a_I^2 + a_R^2)} \quad (14)$$

$$H_U = -\frac{1}{3} \frac{(a_R b_I - a_I b_R)}{(a_I^2 + a_R^2)} \quad (15)$$

It is noted that when viewed from above this point no longer lies on the G -axis and has a general offset in both the G - and H directions. In the special case when the coefficients are purely real, i.e., if $a_I = 0$ and $b_I = 0$, Equations (14) and (15) reduce to $G = -b/3a$ and $H = 0$ as expected.

2. The principal plane of symmetry of surface A no longer coincides with the GA plane. Instead, when viewed from above, this plane is rotated about a normal to the GH plane passing through the now offset umbilical point such that it makes an angle α with lines drawn parallel to the G -axis. It is noted that the overall rotation α of surface A is caused by the presence of one or more of the imaginary parts a_I , b_I and/or c_I of the complex coefficients; only the coefficient $(d_R + id_I)$ has no influence on this rotation.

3. The principal plane of symmetry of surface B is no longer parallel with the HB plane passing through the point $G = -b/3a$. Instead, when viewed from above, this plane is rotated about a normal to the GH plane passing through the now offset umbilical point such that it makes an angle β with lines drawn parallel to the H -axis. Again, the overall rotation β of Surface B is caused by the presence of one or more of the imaginary parts a_I , b_I and/or c_I of the complex coefficients; only the coefficient $(d_R + id_I)$ has no influence on this rotation.
4. Although surfaces A and B are both free to rotate about a normal to the GH plane passing through their common umbilical point, their principal planes of symmetry are 'locked' together such that a fixed right angle is always maintained between them. This infers that the principal planes of symmetry of surfaces A and B both undergo the same angle of rotation in the same direction, i.e., angle $\alpha = \text{angle } \beta$.
5. The vertical location of surface B is no longer fixed at zero—it may vary since Equation (13) now includes a constant term d_I . Hence surfaces A and B , whilst locked in a rotational sense, are free to move independently (i.e., up and down) along their common normal to the GH plane passing through the umbilical point. This allows for more varied intersections—and hence a greater possibility of root combinations—to occur than is possible when compared with the real coefficient case.

Cubic polynomials with generally complex coefficients: Examples

Two examples are now presented which should help illustrate the features just described; the first of these investigates the cubic polynomial with the following complex coefficients:

$$y = (1 - i)x^3 + (-6 - 4i)x^2 + (-9 + 13i)x + (20 + 10i) \quad (16)$$

The roots of this equation are all generally complex and can be shown to be $x = -2 + i$, $1 + 3i$, $2 + i$. The full picture illustrating the location of these roots is found by substituting the various coefficient values from Equation (16) into Equations (12) and (13) to yield the surfaces

$$A = G^3 - 3GH^2 + 3G^2H - H^3 - 6G^2 + 6H^2 + 8GH - 9G - 13H + 20 \quad (17)$$

$$B = 3G^2H - H^3 - G^3 + 3GH^2 - 12GH - 4G^2 + 4H^2 - 9H + 13G + 10 \quad (18)$$

Equations (14) and (15) locate the umbilical point at $G_U = \frac{1}{3}$ and $H_U = \frac{5}{3}$. Figure 11 shows Surface A as a three-dimensional view supplemented by a plan view, the latter clearly indicating the rotation α of the principal plane of symmetry. Figure 12 repeats this information for surface B and, in conjunction with Figure 11, provides visual confirmation of observations (2) and (3) above.

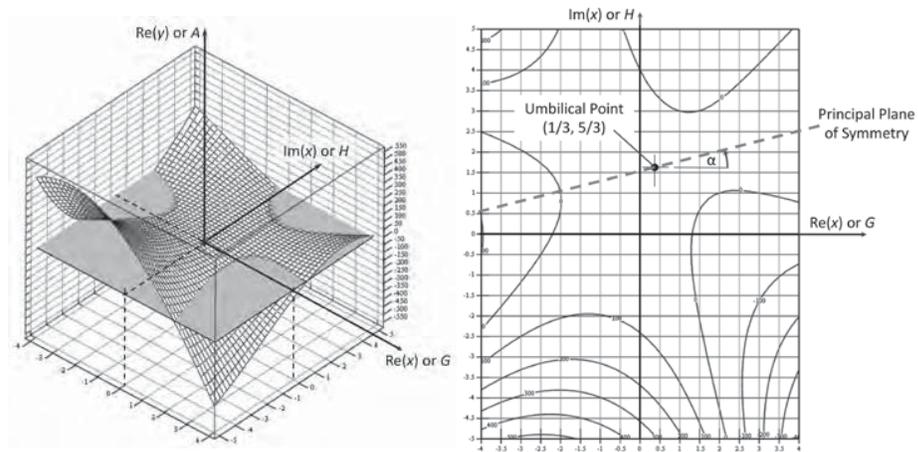


Figure 11. Surface A from Equation (17), showing how the introduction of complex coefficients causes a rotation of the principal plane of symmetry about the umbilical point.

If the plan views of surfaces A and B shown in Figures 11 and 12 respectively were superimposed at the same scale, then it is evident that the principal planes of symmetry (shown by the dashed lines) would cross at right angles, as stated in observation (4). It is further noted that the inclusion of the complex coefficients in the surface description given by Equations (12) and (13) will have a greater scope to influence the overall shape of surfaces A and B when compared with the real coefficient case given by Equations (6) and (7).

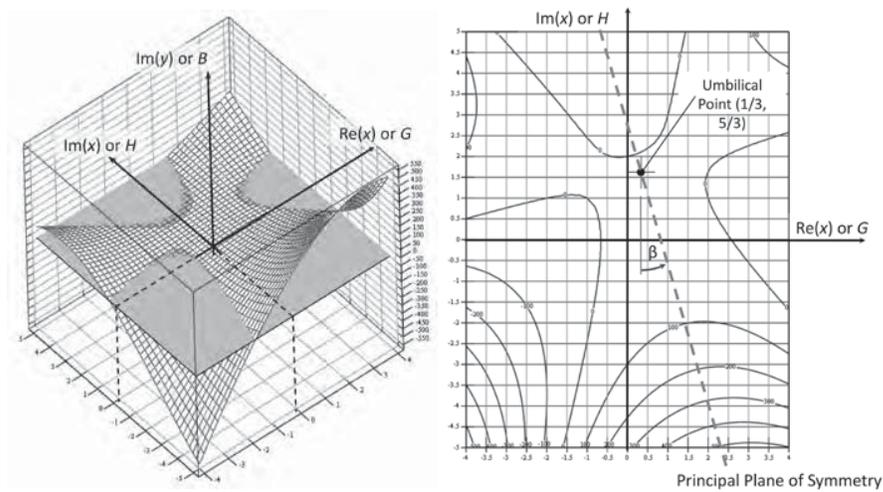


Figure 12. Surface B from Equation (18), showing how the introduction of complex coefficients causes a rotation of the principal plane of symmetry about the umbilical point.

Figure 13 shows the composite of Surfaces A and B and the zero plane, with the roots clearly evident at the three common points of intersection. Figure 14 gives a plan view of the GH (or Argand) plane clearly showing the three complex roots.

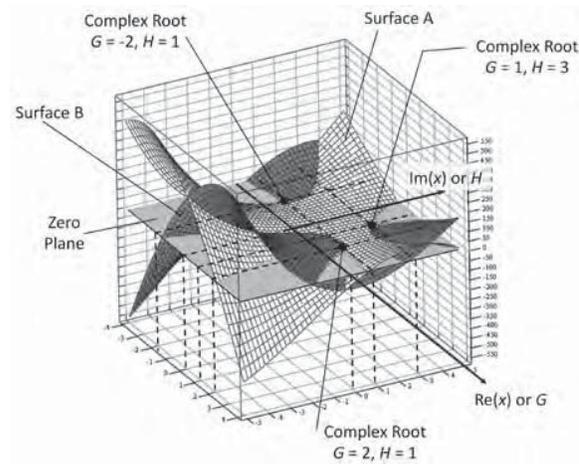


Figure 13. The three roots of the polynomial given by Equation (16) are found from the common intersection of Surface A from Equation (17), Surface B from Equation (18), and the zero plane.

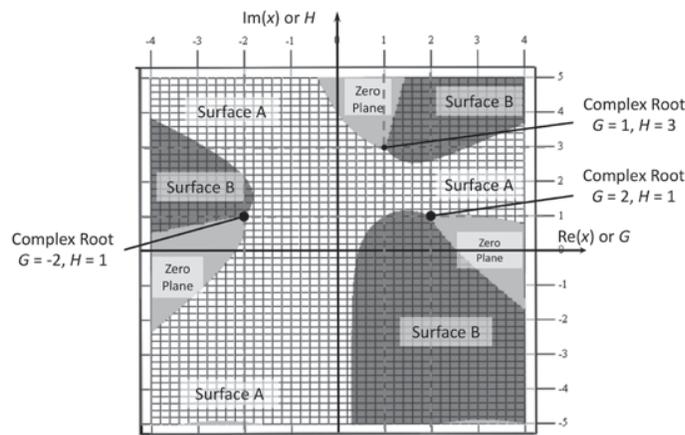


Figure 14. Plan view of Figure 13, clearly showing the three roots of the polynomial given by Equation (16).

It is noted from Figure 14 that the positioning of the three complex roots no longer forms an isosceles triangle in the Argand plane. This observation is generally true for any cubic polynomial with complex coefficients, and is a consequence of there being no necessity for complex conjugate roots to occur. Note however the following special case: if a cubic polynomial with complex coefficients can be rendered in vertex form, the resulting trio of (complex) roots will always define an equilateral triangle.

The second example has been chosen to demonstrate it is possible for a real root to result from a cubic polynomial with complex coefficients.

$$y = (2 + 2i)x^3 + (1 + 3i)x^2 + (2i)x + (1 + i) \tag{19}$$

The three roots of this equation can be shown to be $x = -1, -i, 0.5i$. Substituting the various coefficient values from Equation (19) into Equations (12) and (13) yields the surfaces

$$A = 2G^3 - 6GH^2 - 6G^2H + 2H^3 + G^2 - H^2 - 6GH - 2H + 1 \quad (20)$$

$$B = 6G^2H - 2H^3 + 2G^3 - 6GH^2 + 2GH + 3G^2 - 3H^2 + 2G + 1 \quad (21)$$

Both surfaces share an umbilical point located at $G_U = -\frac{1}{3}$ and $H_U = -\frac{1}{6}$. It is interesting to note that a purely real root can still be found even when the principal plane of symmetry of each surface is no longer parallel with the G- or H-axes. Figure 15 shows a composite view of Surfaces A and B (as given by Equations (20) and (21)) and the zero plane with the roots clearly evident at the three common points of intersection. Figure 16 shows a plan view of Figure 15.

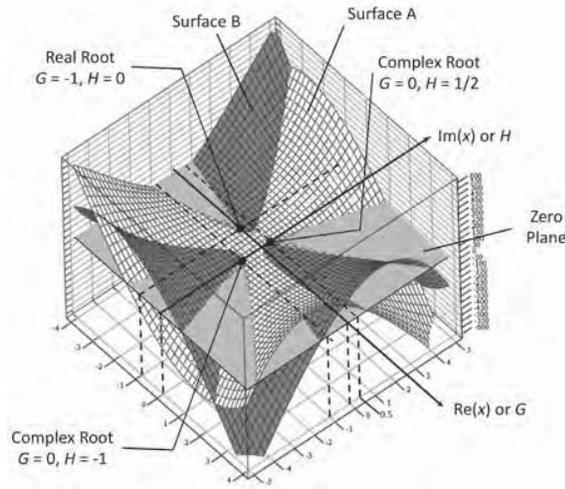


Figure 15. The three roots of the polynomial given by Equation (19) are found from the common intersection of Surface A from Equation (20), Surface B from Equation (21), and the zero plane.

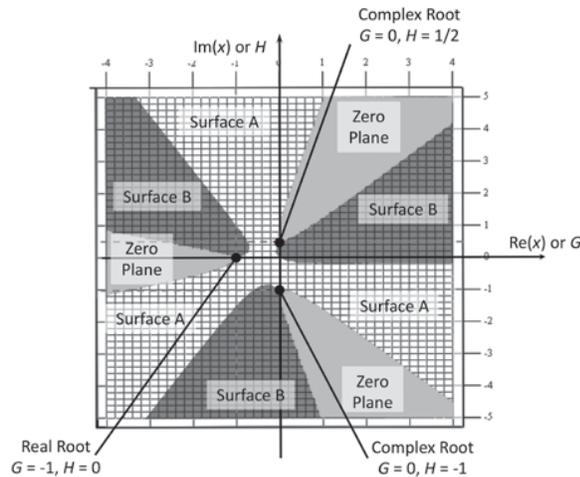


Figure 16. Plan view of Figure 15, clearly showing the three roots of the polynomial given by Equation (19).

Space does not permit the inclusion of the vast number of root arrangements that could occur given the various possible combinations of the complex coefficients. However, it can be stated with certainty that a cubic equation with

generally complex coefficients can yield any possible combination of roots. This is in complete contrast to the real-coefficient counterpart whose roots always occur in predictable arrangements and will always include one real root.

Conclusions

This paper has provided a comprehensive visualization of cubic polynomials with either real or complex coefficients. Whilst the degree of difficulty has increased when compared with the analogous quadratic case (Bardell, 2012, 2014a) it has not detracted from the methodology or the main discoveries reported herein.

This paper has shown that the complex generalisation of a cubic equation with real coefficients may easily be extended to accommodate the case with complex coefficients. The roots are still found from the intersection of the cubic monkey saddle surfaces A and B , representing the real and imaginary parts of the generalisation respectively, with a plane at zero altitude, as first described by Bardell (2012). A cubic equation with generally complex coefficients can yield any possible combination of roots; this is in complete contrast to the real-coefficient counterpart whose roots occur in predictable arrangements. This is primarily due to both monkey saddle surface representations exhibiting a greater freedom of orientation and elevation compared with their real-coefficient counterparts and thus possessing more varied opportunities to intersect each other and the zero plane. A fuller understanding of this, and other phenomena, has been facilitated by the visualization techniques presented in this paper.

Attempts to plot a cubic equation with complex coefficients in only the Cartesian x - y plane fail to reveal any useful information about the location and nature of the roots. Although these roots may be found using various algorithms, it is only possible to visualise them using the methods presented herein.

The simple graphical procedure explained here would also enable a good starting point to be found for a numerical root-finding algorithm, especially in the case when complex coefficients are concerned; the surfaces A and B can easily be computed and plotted from Equations (12) and (13) in most readily available software packages with no need for any complex number manipulations.

The work presented here should help students and teachers alike gain a wider appreciation of cubic polynomials and assist in developing their ability to visualise such concepts as the arrangement of the roots in the three dimensional space framed by the Cartesian and Argand planes. It is anticipated this subject matter could lend itself to an extended assignment using computer algebra systems that would amply satisfy some of the ACARA

(n.d., Rationale) stated aims of Specialist Mathematics, namely “to develop students’ understanding of concepts and techniques drawn from complex numbers”, “ability to solve applied problems using concepts and techniques drawn from complex numbers”, and “capacity to choose and use technology appropriately”.

References

- Abramowitz, M. & Stegun, I. A. (1970). *Handbook of mathematical functions* (9th ed.). New York, NY: Dover Publications.
- Australian Curriculum, Assessment and Reporting Authority. [ACARA]. (n.d.). *Australian Curriculum v. 8 Senior Secondary Curriculum – Mathematical Methods Curriculum*. Unit 1. Retrieved from <http://www.australiancurriculum.edu.au/seniorsecondary/mathematics/mathematical-methods/curriculum/seniorsecondary>
- Australian Curriculum, Assessment and Reporting Authority. (n.d.). *Rationale: Senior Secondary Curriculum – Specialist Mathematics Curriculum*. Retrieved from <http://www.australiancurriculum.edu.au/SeniorSecondary/mathematics/specialist-mathematics/RationaleAims>
- Bardell, N. S. (2012). Visualising the roots of quadratic equations with real coefficients. *Australian Senior Mathematics Journal*, 26(2), 6–20.
- Bardell, N. S. (2014a). Visualising the roots of quadratic equations with complex coefficients. *Australian Senior Mathematics Journal*, 28(1), 6–27.
- Bardell, N. S. (2014b). Some comments on the application of de Moivre’s theorem to solve quadratic equations with real or complex coefficients. *Australian Senior Mathematics Journal*, 28(2), 7–14.
- Board of Studies NSW. (1997). *HSC mathematics extension in NSW*. Retrieved from http://www.boardofstudies.nsw.edu.au/syllabus_hsc/pdf_doc/math4u_syl.pdf
- Cayley, A. (1869). A memoir on cubic surfaces. *Philosophical Transactions*, 159, 231–326.
- Derdzinski, A. (2014). *Math 4552: Cubic equations and Cardano’s formulae*. Retrieved from <https://people.math.osu.edu/derdzinski.1/courses/4552/4552-cubic-quartic.pdf>
- Falconer, K. J. (1986). *The geometry of fractal sets*. Cambridge, UK: Cambridge University Press.
- Harvey Mudd College Department of Mathematics. (n.d.). *The monkey saddle*. Retrieved from https://www.math.hmc.edu/~gu/curves_and_surfaces/surfaces/monkey.html
- Katz, V. J. (1998). *A history of mathematics: An introduction* (3rd ed.). Boston, MA: Addison-Wesley,
- Lipschulz, M. M. (1969). *Schaum’s outline of theory and problems of differential geometry*. New York, NY: McGraw-Hill.
- Mathcad 14.0 M020. (2007). Parametric Technology Corporation.
- Oswald, U. (2009). *The cubic equation*. Retrieved from <http://www.ursoswald.ch/download/CUBIC.pdf>
- Queensland Studies Authority. (2008 amended 2014). *Queensland Mathematics C Senior Syllabus*. Retrieved from https://www.qcaa.qld.edu.au/downloads/senior/snr_maths_c_08_syll.pdf
- Schläfli, L. (1863). On the distribution of surfaces of third order into species, in reference to the absence or presence of singular points, and the reality of their lines. *Philosophical Transactions* 153, 193-241.
- Stroud, K. A. (1986). *Further engineering mathematics*. Basingstoke, UK: Macmillan Education.
- Victorian Curriculum and Assessment Authority [VCAA]. (2010). *VCE Specialist Mathematics*. Retrieved from <http://www.vcaa.vic.edu.au/vce/studies/mathematics/mathsstd.pdf>
- Weisstein, E. W. (n.d.). *Halley’s method*. Retrieved from <http://mathworld.wolfram.com/HalleysMethod.html>