How Newton's Method can be used to find all roots of all polynomials

Catherine Smith-Dance
University of Leicester

April 2020

Abstract

In this paper we will discuss Newton's method, its limitations and a theorem which deals with these limitations. We will be looking in particular at John Hubbard, Dierk Schleicher and Scott Sutherland's theorem on finding all roots of complex polynomials by Newton's method. It will look at constructing a finite set of points so that for every root of every polynomial of fixed degree, at least one of the points will converge to a root under Newton's map.

1 What is Newton's Method?

Newton's method is a way of finding roots of a function. It is an algorithm that is derived from the first few terms of the Taylor series of a function $f(x)$ in the vicinity of a suspected root.

Reminder: Taylor's series of $f(x)$ about the point $x = x_0 + \epsilon$ is given by $f(x_0 + \epsilon) = f(x_0) + f'(x_0)\epsilon + \frac{1}{2}f''(x_0)\epsilon^2 + \cdots$

If we only consider the terms of the first order $f(x_0 + \epsilon) \approx f(x_0) + f'(x_0)\epsilon$ we can estimate the amount of offset ϵ needs to land closer to the root starting from an initial guess x_0. Let $f(x_0 + \epsilon) = 0$ and solve $f(x_0 + \epsilon) \approx f(x_0) + f'(x_0)\epsilon$ for $\epsilon = \epsilon_0$. This gives the equation $\epsilon_0 = -\frac{f(x_0)}{f'(x_0)}$ which is the first order adjustment to the root's position. By letting $x_1 = x_0 + \epsilon_0$ and calculating a new ϵ_1 we can start a process that can be repeated - working out new x_n's and ϵ_n's from previous values - which will eventually lead to it converging to a fixed point (i.e. a root). With a good choice of an initial starting point for the roots position, the algorithm we have just created can be applied iteratively to obtain $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ for $n = 1, 2, 3, ...$

Definition 1.1: If x_n is an approximation of a solution of $f(x) = 0$ and if $f'(x_n) \neq 0$ the next approximation is given by

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

We shall define this as **Newton's Iteration function**, denoted by $N(x) = x - \frac{f(x)}{f'(x)}$. [3]

Example 1.2: Use Newton's method to find all the roots of $x^3 - x^2 - 15x + 1 = 0$ accurate to six decimal places.

Remember the formula for Newton's method: $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$
First of all we need to choose a starting value x_0. One way to establish the interval in which the roots lie is to draw a rough graph of the function so you can see approximately where the roots are.
A way to do this is to draw a table of coordinates between some values of x, allowing you to see roughly where the graph will cross the x-axis and y-axis. If we consider between $x = -5$ and $x = 5$, we get the resulting table of coordinates:

x	-5	-4	-3	-2	-1	0	1	2	3	4	5
f(x)	-74	-19	10	19	14	1	-14	-25	-26	-11	26

As we can see the graph will roughly cross the x axis between $x = -4$ and -3, $x = 0$ and 1, and $x = 4$ and 5.

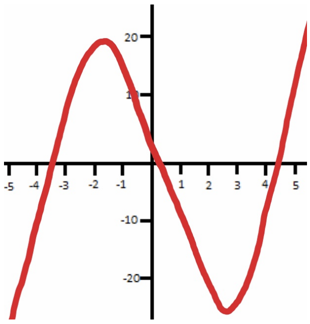

Figure 1.1: Rough graph of $f(x) = x^3 - x^2 - 15x + 1 = 0$

Now we have a rough idea for where the roots lie we can pick an initial x_0 and begin Newton's method to find our first root. We will start with $x_0 = -3.5$.

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = -3.5 - \frac{(-3.5)^3 - (-3.5)^2 - 15 \cdot (-3.5) + 1}{3 \cdot (-3.5)^2 - 2 \cdot (-3.5) - 15} = -3.44347826$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = -3.44347826 - \frac{(-3.44347826)^3 - (-3.44347826)^2 - 15 \cdot (-3.44347826) + 1}{3 \cdot (-3.44347826)^2 - 2 \cdot (-3.44347826) - 15} = -3.44214690$$

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = -3.4421469 - \frac{(-3.4421469)^3 - (-3.4421469)^2 - 15 \cdot (-3.4421469) + 1}{3 \cdot (-3.4421469)^2 - 2 \cdot (-3.4421469) - 15} = -3.4421462$$

As you can see the second and third iterations have now converged to the same point (up to seven decimal places), so we stop here and estimate one of the roots as being $x \approx -3.442146$

We then follow the same procedure to calculate the other two roots with different starting points according to where the graph appears to cross the x-axis.

For $x_0 = 0.5$

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 0.5 - \frac{0.5^3 - 0.5^2 - 15 \cdot 0.5 + 1}{3 \cdot 0.5^2 - 2 \cdot 0.5 - 15} = 0.06557977$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 0.06557977 - \frac{(0.06557977)^3 - (0.06557977)^2 - 15 \cdot 0.06557977 + 1}{3 \cdot (0.06557977)^2 - 2 \cdot 0.06557977 - 15} = 0.06639235$$

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = 0.06639235 - \frac{(0.06639235)^3 - (0.06639235)^2 - 15 \cdot 0.06639235 + 1}{3 \cdot (0.06639235)^2 - 2 \cdot 0.06639235 - 15} = 0.06639231$$

Again, the second and third iterations have converged to the same point (up to seven decimal places) and so we stop the iterations and estimate the second root as $x \approx 0.0663923$.

For $x_0 = 4.5$

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = 4.5 - \frac{4.5^3 - 4.5^2 - 15 \cdot 4.5 + 1}{3 \cdot 4.5^2 - 2 \cdot 4.5 - 15} = 4.38095238$$

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} = 4.38095238 - \frac{(4.38095238)^3 - (4.38095238)^2 - 15 \cdot (4.38095238) + 1}{3 \cdot (4.38095238)^2 - 2 \cdot (4.38095238) - 15} = 4.37575386$$

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = 4.37575386 - \frac{(4.37575386)^3 - (4.37575386)^2 - 15 \cdot (4.37575386) + 1}{3 \cdot (4.37575386)^2 - 2 \cdot (4.37575386) - 15} = 4.37575386$$

Again, the second and third iterations have converged to the same point (up to 8 decimal places) and so again stop the iterations and estimate the final root as $x \approx 4.37575386$.

So the roots of $x^3 - x^2 - 15x + 1 = 0$ accurate to six decimal places are:

$$x \approx -3.442146,\ x \approx 0.066392, \text{ and } x \approx 4.375754$$

2 Problems with Newton's Method

There are three major ways in which Newton's method can fail: when x_0 is the critical point; when there is no root to find; and if the initial starting point (x_0) or iteration coincides with a cycle.

2.1 Critical Point

One of the conditions of Newton's method is $f'(x_n) \neq 0$. So if we were to choose an initial starting point, x_0, such that $f'(x_0) = 0$, then Newton's method will not converge to a root. This can be seen when considering the function $f(x) = x^3 + 1$. If we choose the initial starting point to be $x_0 = 0$ then we can see that $f'(x) = 3x^2 \Rightarrow f'(0) = 0$ and so Newton's method cannot be carried out. This can be illustrated by looking at the graph (Figure 2.1); as you can see the tangent line at $x = 0$ never intersects with the x-axis.

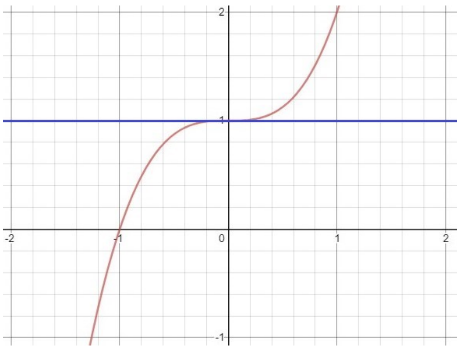

Figure 2.1: Graph of $f(x) = x^3 + 1$ (red) and $x = 1$ (blue)

2.2 No Roots

For Newton's method we are only considering roots in the real plane. So if we are to consider a function such as $f(x) = x^2 + 1$ we will see that it has roots in the complex plane but not in the real plane (shown in the graph, as the function never crosses the x-axis). Therefore in this case Newton's method will not work, as the iterations will never converge to a root.

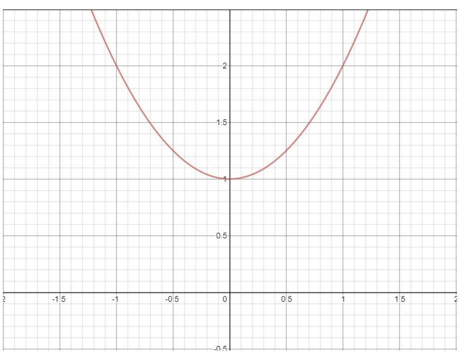

Figure 2.2: Graph of $f(x) = x^2 + 1$

2.3 Periodic Cycle

The third way Newton's method may fail is if the initial starting point or an iteration coincides with a cycle. For example, consider the function $f(x) = x^3 - 2x + 2$ and the initial starting point $x_0 = 1$.

$$x_1 = 1 - \frac{1^3 - 2 \cdot 1 + 2}{3 \cdot 1^2 - 2} = 0; x_2 = 0 - \frac{0^3 - 2 \cdot 0 + 2}{3 \cdot 0^2 - 2} = 1$$

As you can see the iterations have started to cycle, so no matter how many iterations you carry out you will just have alternating answers between 1 and 0, and it will never converge to a root.

3 Newton's Method in the Complex Plane

A solution to one of the problems brought up in Section 2 is to consider Newton's method in the complex plane. Newton's method can be easily generalised to the complex plane, such that for $z = a + bi$, $N(z) = z - \frac{f(z)}{f'(z)}$.

What we have not previously discussed, is that you can view Newton's method as essentially drawing a tangent to the curve and following the tangent to where it crosses the x-axis. If a bad initial starting point is chosen then this means that the tangent to the curve will cross the x-axis a huge distance away from the roots of the function. The way that Newton's method works, will also mean that the second iteration will flip the tangent from one side of the root to the other side. This behaviour of expanding a small area into a large area, results in fractals occurring when considering Newton's method at all points on the curve.

Definition: A **fractal** is a never-ending pattern that repeats itself at different scales.

When considering Newton's method in the complex plane the fractals that appear are really fascinating. Below are some examples of fractals created from Newton's method in the complex plane (called Newton fractals).

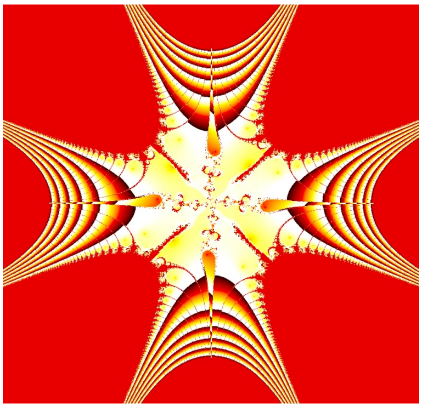

Figure 3.1: Newton's fractal yielded via solving $log(z^4)exp(z^4) - i = 0$ [7]

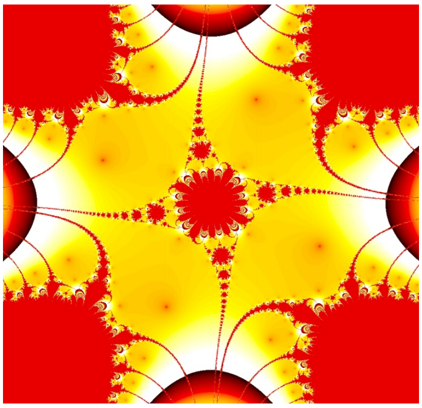

Figure 3.2: Newton's Fractal yielded via solving $exp(z^4) - i = 0$ [7]

As you can see this means that choosing the initial starting point for Newton's method is very important. This brings us nicely onto the theorem of John Hubbard, Dierk Schleicher and Scott Sutherland, on how to find all roots of complex polynomials.

Theorem 3.1: *For every $d \geq 2$ there is a set $\mathcal{S}_d$ consisting of points in $\mathbb{C}$ with the property that for every polynomial $p \in \mathcal{P}_d$ and each of its roots, there is a point $s \in \mathcal{S}_d$ in the basin of the chosen root.* [6]

Before we jump straight into how the theorem works however, we will first discuss a few important properties of Newton's Method that are used in Theorem 3.1.

3.1 Fixed Points

Definition 3.2: A point x_0 is a **fixed point of a function** $f(x)$ if and only if $f(x_0) = x_0$ (i.e. a fixed point is a root of $f(x)$). Moreover, the point x_0 is called an **attracting fixed point** if $|f'(x_0)| < 1$. [1]

An attracting fixed point (x) is when an initial point x_0 is iterated and it always converges to x. Therefore...
Remark 3.3: If a root is an attracting fixed point in Newton's Iteration function, then Newton's method will converge to that root.

Theorem 3.4: *ξ is a root of a function f of multiplicity $k > 0$ if and only if ξ is a fixed point of $N(x)$ (Newton's Iteration function). Moreover, such a fixed point is attracting.* [1]

The above theorem highlights that the fixed points of $N(x)$ are the roots of $f(x)$, and since all fixed points are attracting fixed points when considering Newton's method, we can see that a sequence occurs as you iterate $N(x) : x_0,\ x_1 = N(x_0),\ x_2 = N(x_1), \ldots$ which will converge to the root of $f(x)$.

3.2 Basins of Attraction

When considering functions with multiple roots (e.g. Example 1.2) we notice that when choosing different starting points and carrying out Newton's method, each respective sequence of iterations converges to a different root. Even though the initial starting points in Example 1.2 are relatively close together, their iterations converged to completely different roots; so we must consider which initial points lead to which roots.

Definition 3.5: If ξ is a root of $f(x)$, the **basin of attraction** of ξ is the set of all numbers x_0 such that Newton's method starting at x_0 converges to ξ, i.e. $B(r) = \{x_0 | x_n = N^n(x_0) \text{ converges to } \xi\}$. [10] In other words, the basin of attraction is the points whose orbits attract to x_0.

Definition 3.6: The **immediate basin** of attraction for x_0 is the largest neighbourhood around x_0 contained in the basin of attraction. It can also be considered as the connected component of the basin containing the root.

3.3 Channels to Infinity

Remark 3.7: The point at ∞ is on the boundary of the immediate basin of every root and there are simple arcs in the immediate basin of each root that connect the root to the point at infinity.

> **Definition:** A **simple arc** is simply a curve that is injective, i.e. it does not cross itself.

Definition 3.8: An **access to infinity** is a collection of sets of homotopic paths connecting the root to ∞.

> **Definition:** Two paths with common endpoints are called **homotopic** if one can be continuously deformed into the other leaving the end points fixed and remaining within its defined region.

Theorem 3.9: *If m is the number of critical points of $p(z)$ in the immediate basin of a root ξ, then the immediate basin of ξ has exactly m distinct accesses to infinity.* [6]

Proposition 3.10 *If the number of critical points of the Newton map within some immediate basin is m, counting multiplicities, then this basin has a channel to infinity with width at least $\frac{\pi}{log(m+1)}$. In particular, every basin has a channel with width at least $\frac{\pi}{log(d)}$, where d is the degree of the polynomial.* [6]

With these definitions, remarks and theorems in mind, we can now begin to look at how Theorem 3.1 works. We first of all need to let $\mathcal{P}_d$ be the space of the polynomials of degree d, normalised so that all their roots are in the open unit disk $\mathbb{D}$.

> **Definition:** Let $f(x, y)$ be a polynomial of degree d: $f(x, y) = a_0x^d+a_1x^{d-1}y+a_2x^{d-2}y^2+ ... + a_dy^d + ...$ f is said to be **normalised** if the first non-zero term in the sequence $a_0, a_1, a_2, ..., a_d$ is equal to 1. [4]
>
> **Definition:** The (open) **unit disk** can be considered to be the region in the complex plane defined by $z : |z| < 1$, where $|z|$ is the complex modulus, i.e. it is the interior of a circle of radius 1. [12]

Since all roots are in the open unit disk $\mathbb{D}$, this means that every circle centred at the origin but outside of $\mathbb{D}$ intersects every channel to infinity and hence every such circle intersects the immediate basin of each root. Furthermore, since there is also a lower bound on the width of these channels, this allows for a set $\mathcal{S}_d$ to be created which contains at least one point that will converge to each root, by choosing sufficiently many points spaced along the circle where $\mathcal{S}_d$ is finite.

Henceforth, we can now find all roots of all complex polynomials.

References

[1] A Burton. *Newton's Method and Fractals*. 2009. URL: https://www.whitman.edu/Documents/Academics/Mathematics/burton.pdf. Accessed February 2020.

[2] P. Dawkins. *Newton's Method - Practice Problems Solutions*. 2020. URL: http://tutorial.math.lamar.edu/Solutions/CalcI/NewtonsMethod/Prob5.aspx. Accessed February 2020.

[3] P. Dawkins. *Section 4-13: Newton's Method*. 2018. URL: http://tutorial.math.lamar.edu/Classes/CalcI/NewtonsMethod.aspx. Accessed February 2020.

[4] R. Der and Y. Simsek. "Normalized polynomials and their multiplication formulas". In: *Advances in Difference Equations* 31 (2013). URL: https://advancesindifferenceequations.springeropen.com/articles/10.1186/1687-1847-2013-31.

[5] E. Dummit. *Dynamics, Chaos, Fractals: Introduction to Dynamics*. 2015. URL: https://math.la.asu.edu/~dummit/docs/dynamics_1_introduction_to_dynamics.pdf. Accessed February 2020.

[6] J. Hubbard, D. Schleicher, and S. Sutherland. "How to find all roots of complex polynomials by Newton's method". In: *Inventiones mathematicae* 146 (2001), pp. 1–33.

[7] I. Jackson. *Fractals derived from Newton-Raphson iteration*. 2019. URL: https://www.chiark.greenend.org.uk/~sgtatham/newton/. Accessed April 2020.

[8] A. Jolie. *Newton's method for complex polynomials*. 2015. URL: https://pdfs.semanticscholar.org/138e/fde29c8a84ae58ecc0e5f5d1d076ccc9d14e.pdf. Accessed February 2020.

[9] Osserman. *Homotopy*. 2012. URL: https://www.britannica.com/science/homotopy. Accessed February 2020.

[10] P. D. Straffin. *Newton's Method and Fractal Patterns*. 1991. URL: http://www.zebragraph.com/Polynomiography_files/9971620copy.pdf. Accessed February 2020.

[11] E. W. Weisstein. *Newton's Method*. 2020. URL: http://mathworld.wolfram.com/NewtonsMethod.html. Accessed February 2020.

[12] E. W. Weisstein. *Unit Disk*. 2020. URL: http://mathworld.wolfram.com/UnitDisk.html. Accessed February 2020.

[13] Y. Zhang et al. "From Newton Fractals to Gradient Fractals in Addition to Zhang Fractals". In: *2014 IEEE 17th International Conference on Computational Science and Engineering* (2014), pp. 1237–1242.