Extended Essay in Mathematics

Methods of approximating square roots and the generalization of the method for approximating the $\boldsymbol{n-th}$ root of a positive integer

Name: Luke Minsuk Kim

Candidate Number: 003258-0068

School: South Island School, Hong Kong

Submission Date: October 21 2013

Session: May 2014

Word Count: 3958

Abstract

The research question of the following extended essay is: What are the methods of approximating the value of square roots and how does the efficiency of these methods compare? How can we arrive at a generalization to approximate the value of the $\boldsymbol{n-th}$ root of a positive integer? The essay introduces various methodologies to approximate the value of square roots and compares the number of iterative steps required by each method to converge. Furthermore, certain methods of approximating square roots were generalized to approximate the $n-th$ root of a positive integer.

The essay primarily focused on six different methods of approximating square roots, namely, Babylonian approximation, Bakhshali approximation, Bisection method, Newton's method, approximation using Pell's equation and Iterating linear fractional transformation (ILFT). The methods used in delivering these ideas relied heavily on algebra, iterative sequences and calculus. Other ideas such as matrixes could be applied for concepts such as ILFT to expedite computation. The ideas behind each method were examined and eventually, Bisection, Newton's and ILFT methods could be further generalized to approximate the $n-th$ root. Moreover, the combination of Newton's and Bisection method yielded another mode of $n-th$ root approximation called 'Gould's method'.

Comparison of approximating $\sqrt{k}, k \in \mathbb{Z}^{+}$, comprised of approximating $\sqrt{50}$ and $\sqrt{10^7}$ using different methods. The values obtained were compared to the value displayed on calculator TI-84 (10 and 9 significant figures for $\sqrt{50}$ and $\sqrt{10^7}$ respectively). Results showed that Pell's equation converges the fastest, followed by Newton's, ILFT, Babylonian and Bisection. After the aforementioned methods were generalized to the $n-th$ root, $n=5$ was arbitrarily chosen and their efficiencies were compared. The number of iterative steps to approximate the $n-th$ root for Newton's and Bisection were consistent with those of approximating $\sqrt{k}$ whilst ILFT showed great

disparity. However, more values of n could be examined for better conclusion.

Word Count: 300

Table of Contents

1. Introduction

Surds, which are present in various mathematical calculations ranging from algebra to geometry, are crucial in evaluating values of various quantities. The diagonals of a square and even the golden ratio ($\varphi = \frac{1+\sqrt{5}}{2} = 1.6180339\ldots$) (Wolfram Alpha) involve us to evaluate surds to obtain exact values. A European mathematician called Gherardo of Cremona first used the name 'surds' in the year 1150 (Lawrence). Since then, numerous researches have been conducted and the concept has been put to use in fields such as numerical analysis and discrete mathematics.

Surds are irrational, but non transcendental (notably, $\sqrt{2}$ is a zero of a polynomial $x^2 - 2 = 0$). The following is a famous proof of why surds are irrational numbers. Take $\sqrt{3}$ as an example and we perform our proof by contradiction:

Proof:

$Suppose\ that\ \sqrt{3}\ is\ a\ rational\ number, then\ \sqrt{3}\ can\ be\ represented\ \ as\ the\ following$:

$$\sqrt{3} = \frac{b}{a}, \qquad (1)$$

$$where\ a, b \in \mathbb{Z}, a \neq 0, and\ a, b\ are\ relatively\ prime$$

$$Rearrange\ (1), we\ get$$

$$b = \sqrt{3}\,a \qquad (2)$$

$$Square\ both\ sides\ of\ (2),$$

$$3a^2 = b^2 \qquad (3)$$

$Therefore, \quad b^2\ is\ a\ multiple\ of\ 3, and\ \ b\ is\ also\ a\ multiple\ of\ 3$

$\therefore\ Exists\ k \in \mathbb{Z}, such\ that\ b = 3k\ ,\ substitute\ to\ (3),$

We have

$$3a^2 = (3k)^2$$

$$\therefore\ a^2 = 3k^2 \qquad (4)$$

$Here, a\ must\ be\ a\ multiple\ of\ 3\ since\ a^2\ is\ a\ multiple\ of\ 3.$

$This\ is\ a\ contradictory\ to\ the\ fact\ a\ and\ b\ are\ relatively\ prime.$

$As\ such, \sqrt{3}\ is\ not\ a\ rational\ number => it\ is\ irrational$

Q.E.D.

Since surds are irrational, (i.e. their decimal places go forever without repeating), mathematicians were faced with the question: 'how to approximate the value of surds' long time ago, driven not only by extensive practical applications arisen from architecture and engineering but also by the human desire to accumulate more knowledge and to enjoy mental abstractions. During the past few years, calculators are made to display surds to a certain decimal point (depending on the computer's capacity to store data), and computer programmers had to be informed by mathematicians of what the most efficient algorithm is. More efficient algorithms mean that computers would be able to display satisfactory results more quickly. In addition, computers would only use less processing power to perform calculations, preventing overheating. Also, by inventing various algorithms to approximate surds, computer scientists are able to test computers' speed.

As such, in order to discover the most efficiently working algorithm, mathematicians have experimented with various methods of approximating surds. They initially began by looking at possible methods of approximating square roots, gradually developing those methods to approximate the $n-th$ root of a positive integer. Hence, the title of this exploration is 'what are the methods of approximating the value of square roots and how does the efficiency of these methods compare? How can we arrive at a generalization to approximate the value of the $\boldsymbol{n-th}$ root of a positive **integer?'**

Throughout the exploration, I have two main objectives: First, I will outline the methods that have been developed to approximate the square root of Z^{+}, eventually comparing their efficiency (i.e. how many iterative steps are required for the algorithm to converge?). Second, I will investigate how mathematicians managed to derive an over-arching method for generalizing the $n-th$ root of a positive integer.

Some methods of approximating square roots and surds have not been discussed in the exploration- most notably, the Taylor Series Expansion on a square root function due to its very slow convergence rate compared to most of the known approximation methods.

2. Methods of approximating square roots of a positive value

2.1. The Babylonian approximation

The very first recorded method of approximating square roots comes from the Babylonians, whose calculations were eventually called 'Babylonian algorithm'. The Babylonian algorithm is also commonly associated with Heron of Alexandria, a Greek mathematician, because Heron was the first scholar to explicitly describe the algorithm.

The fundamental idea behind the Babylonian algorithm is that if we have an integer 'n' as an overestimate of the square root of a positive real number 'm', then we can surmise that $\frac{m}{n}$ will be an underestimate. Taking the average of an underestimate and an overestimate would allow us to arrive at a value closer to the actual value of the square root. Below justifies mathematically why the average of the underestimate and the overestimate yields an accurate approximation.

Proof:

$Let\ n \in \mathrm{Z}^{+}, be\ our\ initial\ guess\ of\ the\ square\ root\ of\ a\ number$
$m \in \mathrm{Z}^{+}.$

$(Here, suppose\ n \geq \sqrt{m})$

$Therefore, we\ can\ have$

$$\sqrt{m} = n + E\ , \qquad (5)$$

$$where\ E \in \mathbb{R}, is\ the\ uncertainty, which\ is\ relatively\ small$$

$$Square\ on\ both\ sides,$$

$$m = (n+E)^2 \quad = n^2 + 2nE + E^2 \ \ (6)$$

$Rearranging$ (6),

$$E(E+2n) = m - n^2$$

$$\therefore E = \frac{m-n^2}{E+2n}$$

Since E is a relatively small number,

$$\therefore E \approx \frac{m-n^2}{2n}$$

Therefore, we have $\sqrt{R} = n + E = n + \frac{m-n^2}{2n} = \frac{m+n^2}{2n} = \frac{n+\frac{m}{n}}{2}$

As such, the obtained expression is the average of n and $\frac{m}{n}$

Q.E.D.

The iterative form to approximate $\sqrt{k}, k \in \mathbb{Z}^+$, is $x_{n+1} = \frac{1}{2}(x_n + \frac{k}{x_n})$

Sample Calculation:

Evaluate $\sqrt{146000} \rightarrow We\ take\ an\ overestimate, which\ could\ be\ 400$

$We\ apply\ the\ Babylonian\ algorithm$:

$$First\ iteration, x_1 = \frac{1}{2}\left(400 + \frac{146000}{400}\right) = 382.5$$

$$Second\ iteration, x_2 = \frac{1}{2}\left(382.5 + \frac{146000}{382.5}\right) = 382.0996732$$

$$Third\ iteration, x_3 = \frac{1}{2}\left(382.0996732 + \frac{146000}{382.0996732}\right)$$

$$= 382.0994635$$

We continue the iteration until the value converges for the desired number of decimal places.

Further analysis:

- According to Elgin (30), for $f(x) = x^2 - k, k \in \mathbb{Z}^+$, the simplified iterative form for Newton's method (introduced in page 7) is equivalent to that of the Babylonian approximation. However, this does not imply that the two methods are equivalent as they are based on different mathematical arguments and Newton's method can be generalized to find the $n-th$ root, but it is difficult to do so for Babylonian approximation.

2.2. Bakhshali approximation

Bakhshali approximation, which comes from 17th century Indian mathematicians, provides a suitable approximation of square roots for practical purposes (e.g. engineering & architecture).

The approximation is a more refined version of Heron's formula $\sqrt{R} = \sqrt{(a^2+B)} \approx a + \frac{B}{2a}$, in which $a \in \mathbb{R}$ is the largest approximate to R and $B \in \mathbb{R}$ where $B = R - a^2$ (Archiblad 80). Heron devised this as he studied the Babylonian method. Bakhshali approximation is in the form $\sqrt{R} = a + \frac{B}{2a} - \frac{\left(\frac{B}{2a}\right)^2}{2(a+\frac{B}{2a})}$ - According to Channabasappa (122), the following is how we obtain the expression:

Lemma:

$$R = \left(a + \frac{B}{2a}\right)^2 = a^2 + B + \frac{B^2}{16a^2}$$

$$Assume\ that\ 16a^2 is\ significantly\ larger\ than\ B^2\ \therefore \frac{B^2}{16a^2} \approx 0$$

$$\therefore R \approx a^2 + B, which\ implies\ \sqrt{R} = \sqrt{(a^2+B)} \approx a + \frac{B}{2a}$$

Proof:

$$Let\ \sqrt{R} = \sqrt{a^2+B}\ and\ let\ q_1 = a + \frac{B}{2a}$$

$$(i.e. q_1\ is\ the\ 1st\ approximation\ obtained\ from\ Heron's\ formula)$$

$$By\ substituting\ q_1\ into\ Heron's\ formula\ again, we\ get:$$

$$q_2 = q_1 + \frac{B_1}{2q_1}$$

$$where\ B_1 = R - q_1^2\ and\ q_1 = a + \frac{B}{2a}$$

$$\therefore\ q_2 = q_1 + \frac{R - {q_1}^2}{2q_1}\ which\ is$$

$$q_2 = a + \frac{B}{2a} + \frac{a^2 + B - \left(a + \frac{B}{2a}\right)^2}{2\left(a + \frac{B}{2a}\right)}$$

Upon Simplification, we get the desired expression:

$\sqrt{R} = \sqrt{a^2 + B} = a + \frac{B}{2a} - \frac{\left(\frac{B}{2a}\right)^2}{2(a+\frac{B}{2a})}$

In short, Bakhshali approximation takes the first approximation and substitutes it into Heron's formula to obtain a more accurate approximation.

By taking the third approximation in the form of a second approximation $\left(i.e. q_3 = q_2 + \frac{B_2}{2q_2}\right)$ and by similar substitution methods as above, we can reach an even more refined approximation formula. However, the formula might become too complex to be of practical use. Furthermore, the formula for Bakhshali's approximation is derived from substitution, which makes it difficult to establish an iterative relationship. Thus this approximation is mostly used for rough estimates rather than to get the most accurate value.

Sample Calculation:

Let us take $\sqrt{40}$ then we let R=40, a=6 and B= 4.

$$\sqrt{40} \approx 6 + \frac{1}{3} - \frac{\frac{1}{9}}{\frac{38}{3}} = 6.342105263$$

$$\textit{which is close to the accepted value of } 6.32455532 \ldots$$

2.3. Bisection method

The bisection method is a root finding method, such that the interval in which the root lies is repeatedly bisected in order to find a rough approximation to a solution. For instance, let us take the function $f\ (x) = x^2 - 40$

Note that the function $f(x) = x^2 - 40$ *is continuous in the interval* $[5,7]$ *and* $f(5) < 0, f(7) > 0$

$By\ the\ intermediate\ value\ theorem, f(x)\ must\ contain\ a\ solution\ in\ this\ interval.$

The procedure:

$We\ now\ bisect\ the\ interval$: $\dfrac{5+7}{2} = 6\ and\ f(6) = -4 < 0$

$Since\ f(6) \cdot f(7) < 0, we\ continue\ by\ bisecting\ the\ interval\ [6,7],$

$\frac{6+7}{2} = 6.5\ and\ f(6.5)$ =2.25>0

$Since\ f(6) \cdot f(6.5) < 0,, bisecting\ on\ [6, 6.65] again,$

$\frac{6+6.5}{2} = 6.25\ and\ f(6.25) < 0,$

$f(6.25) \cdot f(6.5) < 0,$

$\therefore\ we\ replace\ 6\ with\ 6.25\ and\ continue\ to\ bisect\ 6.25\ and\ 6.5$

$we\ eventually\ reach\ at\ an\ approximation\ of\ roughly\ \sqrt{40}.$

2.4. The Newton Method

In numerical analysis, Newton's method aims to approximate the roots of real-valued, differentiable functions. Sir Isaac Newton and mathematician Joseph Raphson devised the method.

Firstly, we begin with an initial rough approximation to the root. Then we approximate the function by its tangent line and compute the x-intercept of this tangent line. This x-intercept is a better approximation in majority of cases and we can conduct iterations according to the method aforementioned.

The iterative formula is given by $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$

Proof:

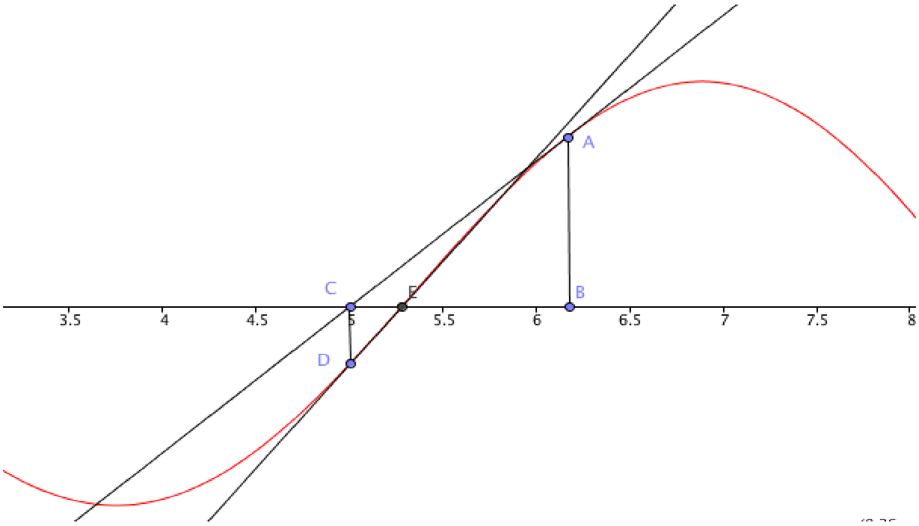

Fig. 1 Demonstration of Newton's method

By Candidate using Geogebra

As Figure 1 shows, our initial guess of the root (x_1), is point A ($x_1, f\ (x_1)$) on the function highlighted as red. Drawing a tangent line of A which meets the x-axis at point C (x_2), which gives us the second point D ($x_2, f\ (x_2)$) on the function. Drawing the tangent line of D allows us to reach at the actual root (zero) of the function. Using this notion, we can develop a formula for the approximation using Newton's method.

$$First, let\ us\ find\ the\ equation\ of\ the\ line\ of\ tanget\ to\ a\ point\ \left(x_1, f(x_1)\right)$$

$$The\ gradient\ would\ be\ f'(x_1) \approx \frac{\Delta y}{\Delta x} = \frac{f(x_1) - 0}{x_1 - x_2}$$

$$Thus, f'(x_1) = \frac{f(x_1)}{x_1 - x_2}, that\ is\ x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

$$Simiarly, if\ we\ get\ our\ third\ x\ intercept\ x_3, f'(x_2) = \frac{f(x_2) - 0}{x_2 - x_3}$$

$$\therefore\ x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}$$

$$Thus\ we\ generalize\ the\ iteration: x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Sample calculation:

$$Take\ f(x) = x^2 - 40, then\ f'(x) = 2x$$

$take\ x_1 = 6\ then\ x_2 = 6 - \frac{36-40}{2*12} = \frac{19}{3} = 6.33333 \ldots..$

Continuing such iteration allows us to arrive at a value that is roughly $\sqrt{40}$

(*We generally take x_1 as a positive integer close to the square root value)

Further analysis:

- Newton's method is not applicable to non-differentiable functions.
- Some curves may fluctuate to an extent that it may yield worse approximations. However, such case is unlikely to happen for functions $f(x) = x^n - k$ $(n \in \mathbb{Z}, k \in \mathbb{Z}^+)$. Therefore, this method can be applied to find the $n - th$ root of a positive integer.

2.5. Approximation using Pell's Equation

Pell's equation, coined after mathematician John Pell, is a Diophantine equation of the form $x^2 - my^2 = 1$, where $x, y, m \in \mathbb{Z}$ and m is a non-perfect square. Using the recursive relationship for Pell's equation allows us to approximate square roots.

Lemma:

$The\ solutions\ for\ Pell's\ equation\ can\ be\ modelled\ by\ the\ relationship$:

$x_{n+1} = x_0 x_n + m y_0 y_n$

$y_{n+1} = y_0 x_n + x_0 y_n$

$where\ x_0\ and\ y_0\ are\ the\ smallest\ integral\ solutions\ to\ Pell's\ equation.$

Proof:

$To\ see\ this, let\ x_{n+1} = ax_n + by_n\ and\ y_{n+1} = cx_n + dy_n, a, b, c, d \in \mathrm{Z}^+$

$(x_{n+1})^2 - m(y_{n+1})^2 = (ax_n + by_n)^2 - m(cx_n + dy_n)^2 = 1$

$= a^2x_n^2 + 2abx_ny_n + b^2y_n^2 - mc^2x_n^2 - 2cdmx_ny_x - md^2y_n^2 = x_n^2 - my_n^2$

$By\ comparing\ like\ terms,$

$2ab - 2cdm = 0 \to ab = cdm\ \therefore a^2b^2 = c^2d^2m^2 \ldots (1)$

$a^2 - mc^2 = 1 \to a^2 = 1 + mc^2 \ldots (2)$

$b^2 - md^2 = -m \to b^2 = md^2 - m \ldots (3)$

$\therefore a^2b^2 = md^2 - m + c^2d^2m^2 - m^2c^2 = c^2d^2m^2$

$\therefore md^2 - m - m^2c^2 = 0 \;\therefore d^2 = 1 + mc^2$

$If\ d = x_0, c = y_0, then\ x_0^2 = 1 + my_0^2\ and\ from\ (2), a = x_0$

$From\ (3),\ b^2 = mx_0^2 - m = m(x_0^2 - 1) = m(my_0^2) = m^2y_0^2$

$\therefore b = my_0$

Combining everything, we finally get

$x_{n+1} = x_0x_n + my_0y_n$

$y_{n+1} = y_0x_n + x_0y_n$

Extension to approximate square root:

$From\ x^2 - my^2 = 1,$

$$we\ can\ have\ \frac{x}{y} = \sqrt{\frac{1}{y^2} + m}\ \ and\ as\ y \to \infty, \frac{x}{y} \to \sqrt{m}$$

Here, we say that $y \to \infty$, because y will become very large in value as we conduct iterations based on the recursive formula obtained.

Sample calculation:

Take $\sqrt{40}$:

$let\ m = 40, x_0 = 19, y_0 = 3\ \ then\ we\ have\ x^2 - 40y^2 = 1$

$Since\ x_{n+1} = 19x_n + 120y_n, and\ y_{n+1} = 3x_n + 19y_n$

$x_1 = 19 * 19 + 120 * 3 = 721\ , y_1 = 3 * 19 + 19 * 3 = 114$

$x_2 = 19 * 721 + 120 * 114 = 27379, y_2 = 3 * 721 + 19 * 114 = 4329$

$$Hence, after\ the\ 2nd\ iteration, we\ have\ \sqrt{40} \approx \frac{27379}{4329} = 6.324555 \ldots$$

2.6. Application of linear transformation for approximation

Linear fractional transformation is a rational function in the form $f(x) = \frac{ax+b}{cx+d}$.

This notion can be developed to approximate square roots. Typically, for $\sqrt{k}, k \in \mathbb{Z}^+, f(x) = \frac{dx+k}{x+d}, where\ d = \lfloor\sqrt{k}\rfloor$. For practicality, we normally begin our iteration by choosing $x_1 = d$.

Therefore, the iterative form to approximate $\sqrt{k}\ is\ x_n = f(x_{n-1}) = \frac{dx_{n-1}+k}{x_{n-1}+d}$, where $d = \lfloor\sqrt{k}\rfloor$

Proof:

$$Let\ x = \frac{dx+k}{x+d}\ and\ we\ get\ x^2 = k\ \therefore x = \sqrt{k}\ (x, k > 0)$$

Thus the linear transformation gives us a valid method of approximating square roots.

Sample calculation:

Using $\sqrt{40}$ for demonstration, we can write:

$$f(x) = \frac{6x+40}{x+6}\ as\ a\ linear\ fractional\ transformation$$

Now, in order to get our approximated value more quickly, we iterate $f\ (x)$ by x_1=6. Then

x_2= $f(6) = \frac{36+40}{6+6} = \frac{19}{3} = 6.3333333 \ldots$

x_3= $f\ (x_2)$=$f\big(f(6)\big) = \frac{38+40}{\frac{19}{3}+6} = 6.3243 \ldots$

Continuing this, we see that the iteration approaches the value of $\sqrt{40}$ very quickly.

Figure 2 shows $f\ (x)$=$\frac{6x+40}{x+6}$ and $y = x$. Notice that $y = x$ will intersect $f\ (x)$ at exactly $(\sqrt{40}, \sqrt{40})$. When we have $x = 6$, we have B, and then we compute $ff(6)$ to get C. As we carry on, our point gets closer to the intersection $A\big(\sqrt{40}, \sqrt{40}\big)$. Thus it can show us that as we increase the number of our iterations, the value will get closer to the intersection.

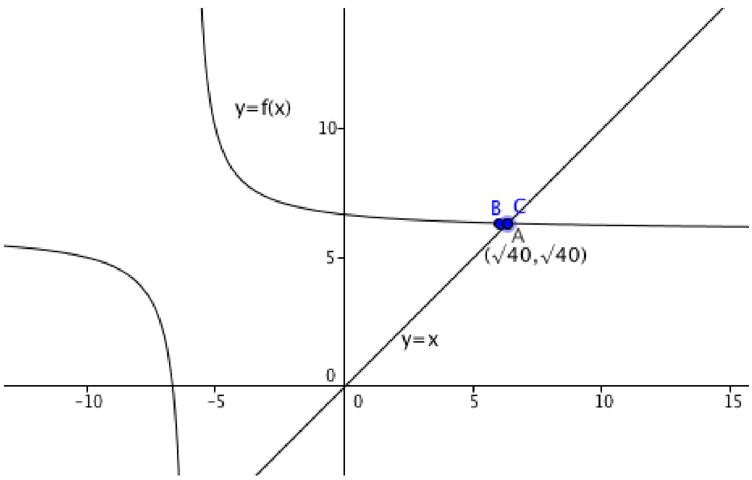

Fig. 2. Graph of $f\ (x)$ and $y = x$ (O'dorney)

To expedite the process, the definition of *Mobius* Transformation can be used. For matrix $\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we define the product $\alpha \cdot x = \frac{ax+b}{cx+d}$ for a point x on the Riemann Sphere $(the\ extended\ complex\ plane\ \bar{C} = C \cup \{\infty\})$ (Thorup 5)

This transformation $x \rightarrow \alpha \cdot x$ of the Riemann sphere is the Mobius transformation. This means that iterating for the $n-th$ time is equivalent to the product $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^n \cdot x$

Sample calculation:

$$\begin{pmatrix} 6 & 40 \\ 1 & 6 \end{pmatrix} \cdot x = \frac{6x+40}{x+6}$$

$$\begin{pmatrix} 6 & 40 \\ 1 & 6 \end{pmatrix}^4 = \begin{pmatrix} 11536 & 72960 \\ 1824 & 11536 \end{pmatrix}$$

$$we\ define\ f(x) = \begin{pmatrix} 11536 & 72960 \\ 1824 & 11536 \end{pmatrix} \cdot x = \frac{11536x+72960}{1824x+11536}.$$

$Now, f(6) = 6.32455516 \ldots (1)$

The value of (1) is the value we would obtain after 4 iterations.

This method is potentially quicker for computers to evaluate, as they mostly need to carry out matrix multiplications.

2.7. Comparison

In this section, I will compare different methods aforementioned in order to compare how quickly the approximations converge. For the approximation, one relatively small number ($\sqrt{50}$) and one relatively big number ($\sqrt{10^7}$) will be used.

The approximations will be computed to 10 significant figures for $\sqrt{50}$ and 9 significant figures for $\sqrt{10^7}$ and will be compared with the accepted value – the number represented by a calculator (provided below). The error is computed by subtracting the approximation from the accepted value.

Convergence is thought to be reached when the approximate value and the error stays constant for 2 consecutive iterations.

(*Note:* The Bakhshali approximation will be excluded, as it is not perfectly suited for iteration.)

(From hand held calculator - TI-84)

The accepted value for $\sqrt{50}$ is 7.071067812 and for $\sqrt{10^7}$, it is 3162.27766

Babylonian method:

$$\sqrt{50}$$

We take an arbitrary (but reasonable) overestimate, which would be 10

No.of iterations	Approximate value	Error
1	7.5	0.428932188
2	7.083333333	0.012265521
3	7.071078431	0.000010619
4	7.071067812	0
5	7.071067812	0

$$\sqrt{10^7}$$

We take an overestimate, which would be 3500

No.of iterations	Approximate value	Error
1	3178.571429	16.29376857
2	3162.319422	0.041762151
3	3162.27766	0.000000444
4	3162.27766	0.000000168
5	3162.27766	0.000000168

Newton's Method:

$$\sqrt{50}$$

We take $f(x) = x^2 - 50 \rightarrow f'(x) = 2x$

No.of iterations	Approximate Value	Error
1	7.071428571	0.00036075
2	7.071067821	0
3	7.071067812	-0.000000009
4	7.071067812	-0.000000009

$$\sqrt{10^7}$$

We take $f(x) = x^2 - 10^7 \rightarrow f'(x) = 2x$

No.of iterations	Approximate value	Error
1	3162.277918	0.00025838
2	3162.27766	0.000000168
3	3162.27766	0.000000168

Bisection Method:

$$\sqrt{50}$$

$$We\ see\ that\ the\ funciton\ f(x) = x^2 - 50\ is\ continuous\ in\ the\ interval\ [6,8]\ and\ f(6) < 0, f(8) > 0.$$

(* a, b are the values of the Min and Max of the interval, $c = \frac{a+b}{2}$)

No.of iterations	value of a	value of b	value of c	f(c)	Error
1	6	8	7	-1	-0.071067812
2	7	8	7.5	6.25	0.428932188
3	7	7.5	7.25	2.5625	0.178932188

4	7	7.25	7.125	0.765625	0.053932188
5	7	7.125	7.0625	-0.12109375	-0.008567812
6	7.0625	7.125	7.09375	0.321289063	0.022682188
7	7.0625	7.09375	7.078125	0.099853516	0.007057188
8	7.0625	7.078125	7.0703125	-0.010681152	-0.000755312
9	7.0703125	7.078125	7.07421875	0.044570923	0.003150938
10	7.0703125	7.07421875	7.072265625	0.016941071	0.001197813
11	7.0703125	7.072265625	7.071289063	0.003129005	0.000221251
12	7.0703125	7.071289063	7.070800781	-0.003776312	-0.000267031
13	7.070800781	7.071289063	7.071044922	-0.000323713	-0.00002289
14	7.071044922	7.071289063	7.071166992	0.001402631	0.00009918
15	7.071044922	7.071166992	7.071105957	0.000539456	0.000038145
16	7.071044922	7.071105957	7.071075439	0.00010787	0.000007627
17	7.071044922	7.071075439	7.071060181	-0.000107921	-0.000007631
18	7.071060181	7.071075439	7.07106781	-2.55532E-08	-0.000000002
19	7.07106781	7.071075439	7.071071625	5.39224E-05	0.000003813
20	7.07106781	7.071071625	7.071069717	2.69484E-05	0.000001905
21	7.07106781	7.071069717	7.071068764	1.34614E-05	0.000000952
22	7.07106781	7.071068764	7.071068287	6.71794E-06	0.000000475
23	7.07106781	7.071068287	7.071068048	3.34619E-06	0.000000236
24	7.07106781	7.071068048	7.071067929	1.66032E-06	0.000000117
25	7.07106781	7.071067929	7.07106787	8.17384E-07	0.000000058
26	7.07106781	7.07106787	7.07106784	3.95915E-07	0.000000028
27	7.07106781	7.07106784	7.071067825	1.85181E-07	0.000000013
28	7.07106781	7.071067825	7.071067818	7.9814E-08	0.000000006
29	7.07106781	7.071067818	7.071067814	2.71304E-08	0.000000002
30	7.07106781	7.071067814	7.071067812	7.88617E-10	0
31	7.07106781	7.071067812	7.071067811	-1.23823E-08	-0.000000001
32	7.071067811	7.071067812	7.071067811	-5.79683E-09	-0.000000001
33	7.071067811	7.071067812	7.071067812	-2.5041E-09	0
34	7.071067812	7.071067812	7.071067812	-8.57746E-10	0

$$\sqrt{10^7}$$

$We\ see\ that\ function\ f(x) = x^2 - 10^7\ is\ continuous\ in\ the\ interval$ $[3161,3163],\ f(3161) < 0\ and\ f(3163) > 0.$

No.of iterations	Value of a	Value of b	Value of c	f(c)	Error
1	3161	3163	3162	-1756	-0.27766
2	3162	3163	3162.5	1406.25	0.22234
3	3162	3162.5	3162.25	-174.9375	-0.02766
4	3162.25	3162.5	3162.375	615.640625	0.09734
5	3162.25	3162.375	3162.3125	220.3476563	0.03484
6	3162.25	3162.3125	3162.28125	22.70410156	0.00359
7	3162.25	3162.28125	3162.265625	-76.11694336	-0.012035
8	3162.265625	3162.28125	3162.273438	-26.70648193	-0.0042225
9	3162.273438	3162.28125	3162.277344	-2.001205444	-0.00031625
10	3162.277344	3162.28125	3162.279297	10.35144424	0.001636875
11	3162.277344	3162.279297	3162.27832	4.175118446	0.000660312
12	3162.277344	3162.27832	3162.277832	1.086956263	0.000172031
13	3162.277344	3162.277832	3162.277588	-0.45712465	-0.000072109
14	3162.277588	3162.277832	3162.27771	0.314915791	0.000049961
15	3162.277588	3162.27771	3162.277649	-0.071104433	-0.000011074
16	3162.277649	3162.27771	3162.277679	0.121905677	0.000019443
17	3162.277649	3162.277679	3162.277664	0.025400622	0.000004185
18	3162.277649	3162.277664	3162.277657	-0.022851905	-0.000003445
19	3162.277657	3162.277664	3162.27766	0.001274358	0.00000037
20	3162.277657	3162.27766	3162.277658	-0.010788774	-0.000001537
21	3162.277658	3162.27766	3162.277659	-0.004757209	-0.000000584
22	3162.277659	3162.27766	3162.27766	-0.001741424	-0.000000107
23	3162.27766	3162.27766	3162.27766	-0.000233533	0.000000131
24	3162.27766	3162.27766	3162.27766	0.000520412	0.000000251
25	3162.27766	3162.27766	3162.27766	0.00014344	0.000000191
26	3162.27766	3162.27766	3162.27766	-4.50462E-05	0.000000161
27	3162.27766	3162.27766	3162.27766	4.91962E-05	0.000000176

28	3162.27766	3162.27766	3162.27766	2.07499E-06	0.000000169
29	3162.27766	3162.27766	3162.27766	-2.14856E-05	0.000000165
30	3162.27766	3162.27766	3162.27766	-9.70624E-06	0.000000167
31	3162.27766	3162.27766	3162.27766	-3.8147E-06	0.000000168
32	3162.27766	3162.27766	3162.27766	-8.69855E-07	0.000000168

Iterating Linear Fraction:

$$\sqrt{50}$$

Take $f(x) = \frac{7x+50}{x+7}$ $and\ start\ from\ x = 7$

No.of iterations	Approximate value	Error
1	7.071428571	0.000360759
2	7.07106599	-0.000001822
3	7.071067821	0.000000009
4	7.071067812	0
5	7.071067812	0

$$\sqrt{10^7}$$

Take $f(x) = \frac{3162x+10^7}{x+3162}$ $and\ start\ from\ x = 3162$

No.of iterations	Approximate value	Error
1	3162.277672	0.000012359
2	3162.27766	0.000000168
3	3162.27766	0.000000168

Pell's Equation:

$$\sqrt{50}$$

$$initial\ solution\ x = 99, y = 14$$

$$Recursion: x_{n+1} = 99x_n + 700y_n, y_{n+1} = 14x_n + 99y_n$$

(* a, b are initial values and c, d are values obtained after iteration)

No.of iterations	a	b	c	d	Approx. Value	Error
1	99	14	19601	2772	7.071067821	0.000000009
2	19601	2772	3880899	548842	7.071067812	0
3	3880899	548842	768398401	108667944	7.071067812	0

$$\sqrt{10^3}$$

$$initial\ solution\ x = 39480499, y = 1248483$$

$$Recursion: x_{n+1} = 3948049x_n + 1248483000y_n,$$

$$y_{n+1} = 1248483x_n + 39480499y_n$$

(* Reason for using $\sqrt{10^3}$ explained in Evaluation)

No.of iterations	a	b	c	D	Approx. Value	Error
1	39480499	1248483	3.11742E+15	9.85815E+13	31.6227766	0.000000002
2	3.11742E+15	9.85815E+13	2.46155E+23	7.78409E+21	31.6227766	0.000000002

Conclusion:

Summary of convergence

For approximating $\sqrt{50}$:

- Babylonian: 4 steps, error $(e) = 0$
- Newton's method: 3 steps, $e = -0.000000009$
- Bisection method: 33 steps, $e = 0$
- Iterating Linear Fraction: 4 steps, $e = 0$
- Pell's equation: 2 steps, $e = 0$

For approximating $\sqrt{10^7}$:

- Babylonian: 4 steps, $e = 0.000000168$
- Newton's method: 2 steps, $e = 0.000000168$
- Bisection method: 31 steps, $e = 0.000000168$
- Iterating Linear Fraction: 2 steps, $e = 0.000000168$

For approximating $\sqrt{10^3}$:

- Pell's equation: 1 step, $e = 0.000000002$

Analysis:

- Pell's equation is very powerful, but we have to spend time to find the initial solution that we can use for iteration. The process of finding x_0 *and* y_0 is difficult for large numbers.
- Newton's method, Iterating Linear Fraction, and Babylonian method follow in terms of efficiency.
- Bisection method converges the slowest, but with this procedure, we are certain of finding at least one root.

Evaluation:

- There is not much difference in terms of convergence for a relatively big and small number
- 10^3 was used for Pell's equation because the initial integral solutions for $x^2 - 10^7y^2 = 1$ is too large to be entered into Microsoft Excel.
- Excel's *floating point*[1] error might make the value of the error in the table slightly inaccurate, but it nonetheless gives us an insight into convergence.
- As the number of significant figures instead of decimal places was used, the number of decimal places explored for a big number was much less than that of a smaller number. This might explain why the number of iterative steps for a bigger number is less than that of a smaller number.
- The Babylonian approximation may have converged 1~2 steps quicker if the initial overestimate is very close to the actual value of the square root.

[1] (Computing) a method of representing approximations of real numbers

Graphical analysis: value against the number of iterations

$\sqrt{50}$

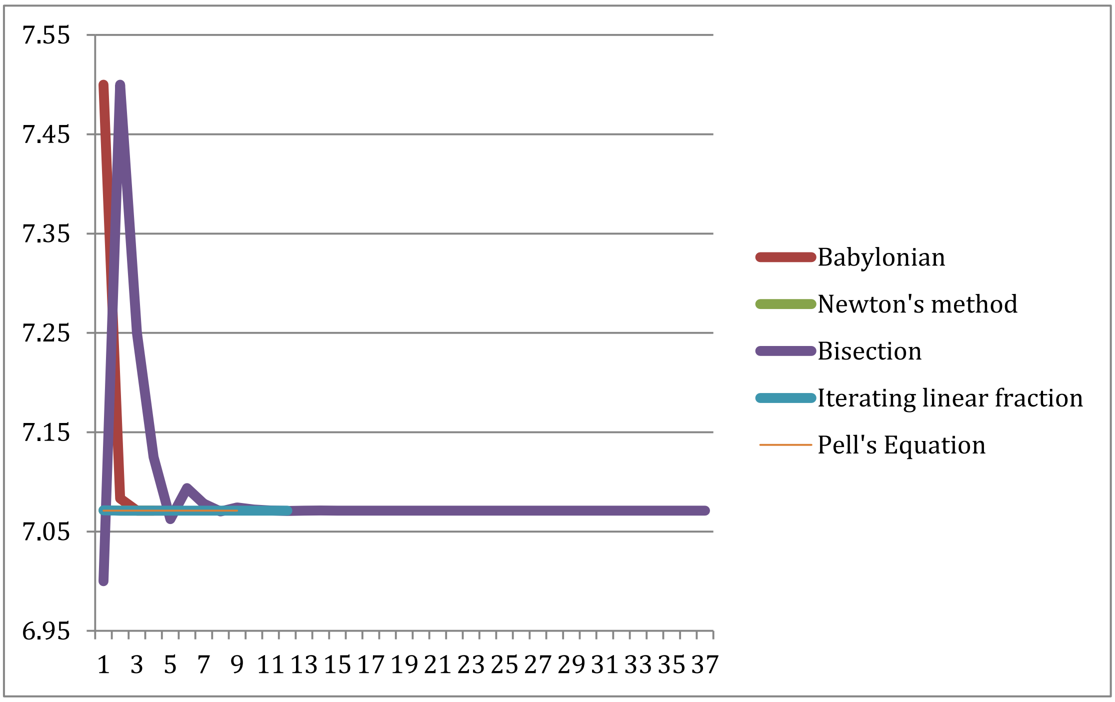

Fig.3. By Candidate using Excel

The Babylonian method drops down very quickly from the overestimate to converge. The bisection method shows 2 major oscillations from the convergence value before converging entirely. Others converge very quickly from the first iteration.

$\sqrt{10^7}$

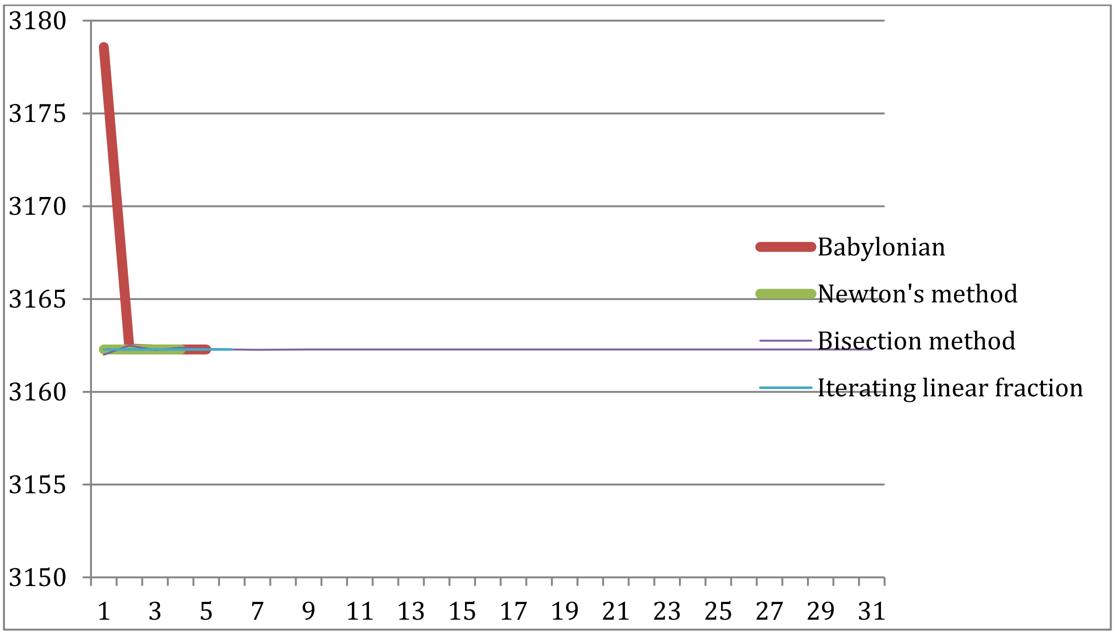

Fig.4. By Candidate using Excel

For clarity, we get rid of the Babylonian, which has a large initial value.

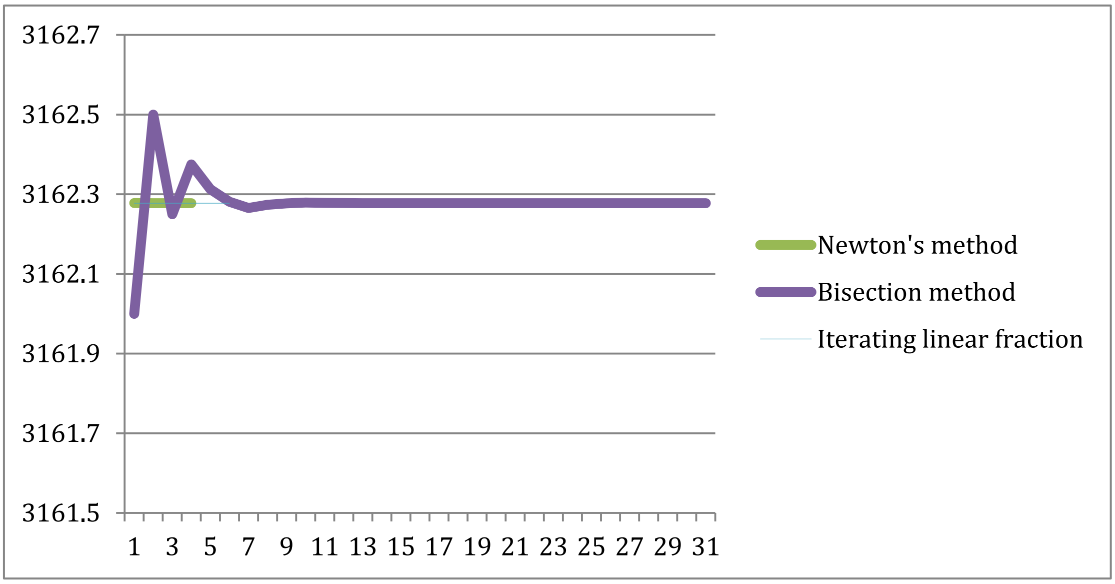

Fig.5. By Candidate using Excel

As with $\sqrt{50}$, the bisection method oscillates, whereas the others converge quickly.

3. Generalization to approximate the value of an $\boldsymbol{n-th}$ root of a number

3.1. Bisection method

As a natural extension, we take function $f(x) = x^n - k, k \in \mathbb{Z}^+$ to approximate the $n - th$ root of k.

$$Find\ a, b\ such\ that\ a^n < x^n < b^n \rightarrow a, b\ is\ our\ initial\ interval$$

$$a, b, \in \mathbb{Z}^+, f(a) < 0, f(b) > 0, let\ c = \frac{a+b}{2}$$

$$if\ f(c) = 0, we\ stop, otherwise\ if\ f(c)f(a) < 0,$$

$$our\ new\ interval\ is\ (a, c)$$

$$if\ f(c)f(a) > 0, our\ new\ interval\ is\ (c, b)$$

Sample calculation:

Approximate the $5th\ root\ of\ 40$.

$$f(x) = x^5 - 40\ is\ continuous\ in\ the\ interval\ [1,3], f(1) < 0, f(3) > 0$$

$$\frac{1+3}{2} = 2\ and\ f(2) < 0 \therefore take\ 2,3$$

$$\frac{2+3}{2} = 2.5\ and\ f(2.5) > 0\ \therefore take\ 2,2.5$$

Continue until we get a convergence $\sqrt[5]{40} \approx 2.091279105$

3.2. Newton's method

For a function $f(x) = x^n - k,\ f'(x) = nx^{n-1}$

We then form an iterative relationship:

$$x_{m+1} = x_m - \frac{x_m^n - k}{n * x_m^{n-1}}$$

Sample calculation:

$$let\ f(x) = x^5 - 40$$

$$f'(x) = 5x^4$$

$$take\ x_1 = 2\ \therefore x_2 = 2 - \frac{2^5 - 40}{5 * 2^4} = 2.1$$

Continue until convergence.

3.3. Generalized linear transformation

$$For\ \sqrt{k}, f(x) = \frac{dx + k}{x + d}\ was\ our\ linear\ fraction, because$$

$$f(x) = x\ yields\ x = \sqrt{k}$$

$$For\ \sqrt[3]{k}, f(x) = \frac{d(x + x^2) + k}{x^2 + d(1 + x)} = x$$

$$\therefore x^3 + dx(1 + x) = dx + dx^2 + k$$

$$\therefore x = \sqrt[3]{k}$$

In general, for $\sqrt[n]{k}, we\ iterate\ f(x) = \frac{d\sum_{i=1}^{n-1} x^i + k}{x^{n-1} + d\sum_{i=0}^{n-2} x^i}$

To check, $\frac{d\sum_{i=1}^{n-1} x^i + k}{x^{n-1} + d\sum_{i=0}^{n-2} x^i} = x$

$$\therefore x^n + dx(1 + x + x^2 + \cdots + x^{n-2}) = d(x + x^2 + \cdots + x^{n-1}) + k$$

$$\therefore x = \sqrt[n]{k}$$

The iterative form $x_m = f(x_{m-1}) = \frac{d\sum_{i=1}^{n-1} x_{m-1}^i + k}{x_{m-1}^{n-1} + d\sum_{i=0}^{n-2} x_{m-1}^i}, d = \lfloor \sqrt{k} \rfloor$

Sample calculation:

Approximate the $5th\ root\ of\ 40, then$

$$f(x) = \frac{2(x + x^2 + x^3 + x^4) + 40}{x^3 + 2(1 + x + x^2 + x^3)}$$

$$if\ x = 2, f(2) = \frac{2(2 + 2^2 + 2^3 + 2^4) + 40}{2^3 + 2(1 + 2 + 2^2 + 2^3)} \approx 2.63157 \ldots$$

Continue until we reach convergence.

Further analysis:

- When we generalize to the $n - th$ root, we cannot call this method iterating linear fraction, as it becomes a polynomial.
- Mobius transformation cannot be applied to $f\ (x)$ that is in non-linear form. Thus matrix calculations cannot replace iterations.

3.4. Gould's method

Gould's method is a further extension of the bisection method and the Newton's method, which tries to find a root of a continuous and differentiable function $f(x)\ in\ [x_1, x_2]\ where\ f(x_1) \cdot f(x_2) < 0$

$i.e., f(x)$ $must\ pass\ through\ at\ least\ one\ zero\ in\ [x_1, x_2].$

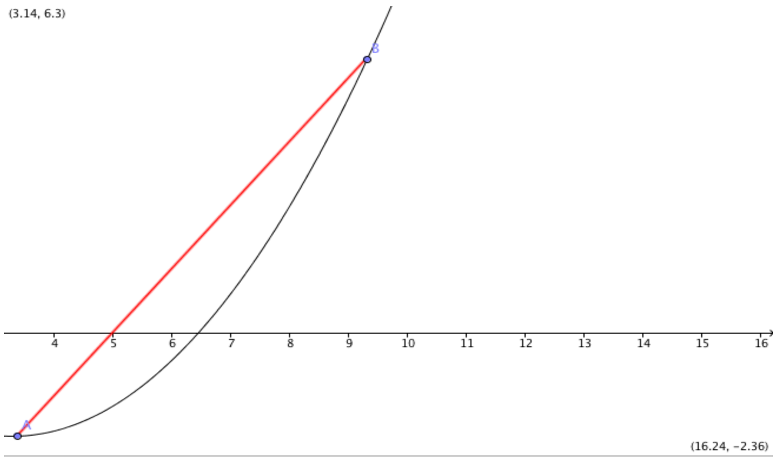

Fig.6. (Gould 65)

Imagine that there is a function $f\ (x)$. We place two points, A and B with x coordinates x_1 and x_2. We connect them by a line. Note that we have an assumption here, since we believe that $f\ (x)$ is similar to a linear function.

Now, the x intercept of the red line on the graph is close to the value of the root of the function $f\ (x) = x^n - k$.

By setting up a relationship for the gradient,

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} \approx f'(x_2)$$

$$x_3\ is\ our\ new\ approximation$$

$$\therefore\ x_3 = x_2 - \frac{f(x_2)}{f'(x_2)} = x_2 - \frac{x_2^n - k}{\frac{x_1^n - x_2^n}{x_1 - x_2}} = x_2 - \frac{x_2^n - k}{\sum_{i=0}^{n-1} x_1^i x_2^{n-1-i}}$$

$$If\ f(x_3)f(x_2) < 0, we\ replace\ x_1\ with\ x_3, else\ we\ replace\ x_2\ with\ x_3$$

$$Eventually\ we\ get\ an\ iteration: x_n = x_{n-1} - \frac{x_{n-1}^2 - k}{\sum_{i=0}^{n-1} x_{n-2}^i x_{n-1}^{n-1-i}}$$

As the graph below illustrates, we get closer to the real root after a number of approximations.

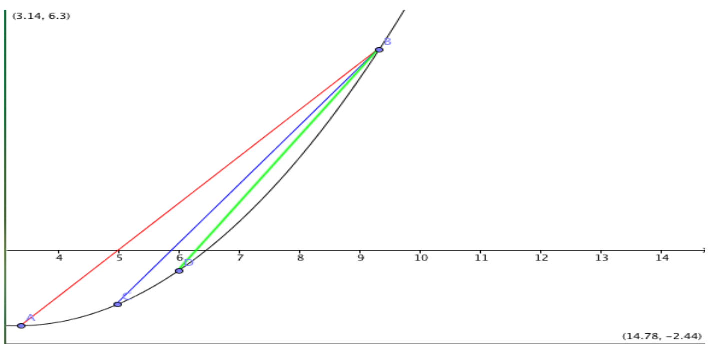

Fig.7. (Gould 66)

Sample calculation:

$$f(x) = x^5 - 40$$

$$let\ x_1 = 1, x_2 = 3, f(x_1) < 0, f(x_2) > 0\ then\ x_3 = 3 - \frac{3^5 - 40}{\sum_{i=0}^{4} 1^i 3^{4-i}}$$

$$= 1.3223 \ldots$$

$$f(x_3) < 0\ \therefore continue\ the\ iteration\ with\ x_2\ and\ x_3$$

3.5. Comparison:

From comparing square root values, we learnt that the size of the number does not greatly affect the rate of convergence as long as the method used is the same. We assume the case is the same for the $n-th$ root. Thus we only evaluate $\sqrt[5]{50}$ to 10 significant figures and compare with the value displayed on TI-84. The fifth root is arbitrarily chosen for comparison.

(As Gould's method has not been introduced for approximating square roots, it won't be included in the comparison for convergence)

The accepted value for $\sqrt[5]{50}$ is 2.186724148

Conclusion:

Summary of convergence:

- Bisection: 32 steps, $e = 0$
- Newton: 4 steps, $e = 0$
- Iterating linear fraction: 225 steps, $e = 0$

Analysis:

- The number of iterative steps needed to approximate the $n-th$ root for Newton's method and bisection method is consistent with approximating square root values.
- Upon extension to approximate a $n-th$ root, the generalized linear fractional transformation requires an excessive number of steps to converge. It is possible that the approximation converges slower as n

increases due to the increasing number of expressions in a polynomial.

Graphical Analysis:

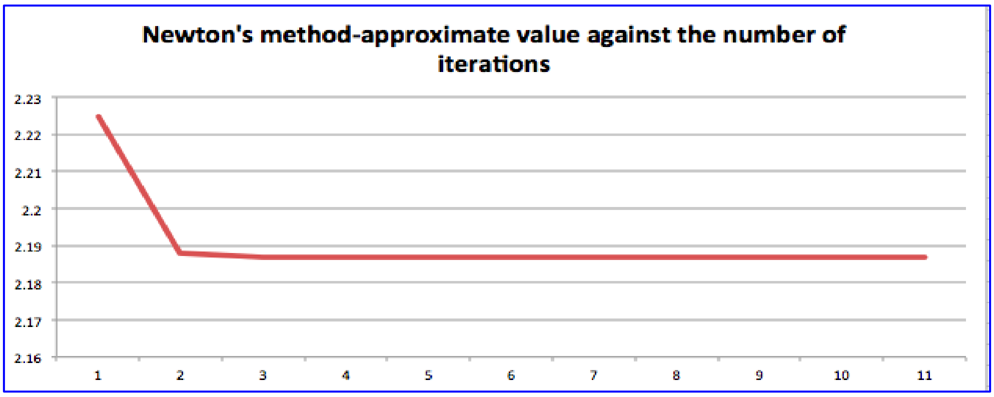

Fig.8. By Candidate using Excel

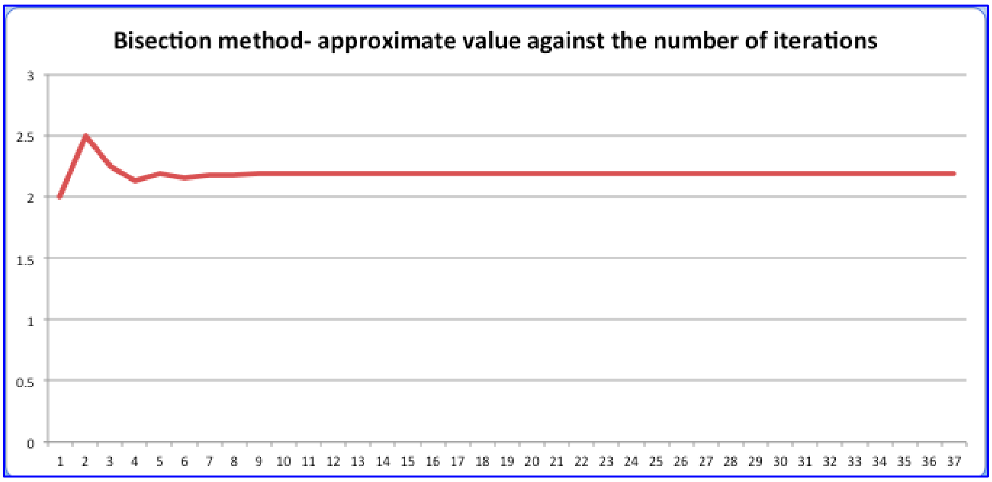

Fig.9. By Candidate using Excel

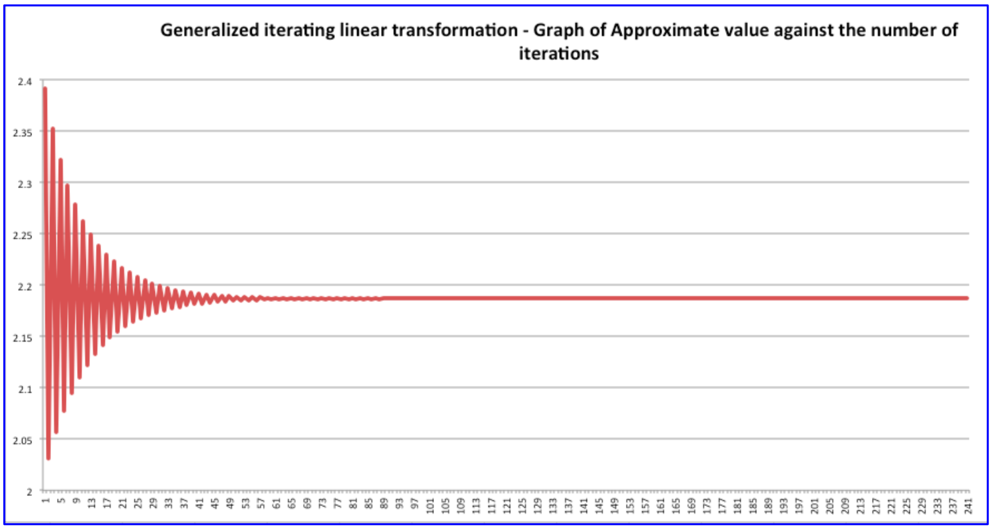

Fig.10. By Candidate using Excel

Newton's method quickly converges as it drops from the initial approximation that is relatively large. There are minor oscillations for the bisection method, but it nonetheless quickly converges. The extreme oscillations in value and slow convergence are evident for the generalized iterating linear fractional transformation.

Evaluation:

- $n = 6\ and\ 7$ have also been tested. Newton's and Bisection stayed within the range of less than 10 and less than 40 iterations respectively. The number of iterations increased for generalized linear fraction. However, it is difficult to determine the efficiency of an algorithm for approximating the $n - th$ root by only investigating certain n. More values could be examined for a much more general trend.
- It is unclear whether the size of the number k for $\sqrt[n]{k}$ affects convergence.

4. Conclusion

In sum, the investigation showed that Pell's equation is the most efficient method to approximate square root. However, it contains certain restrictions, which makes Newton's method more preferable. Newton's method also proved to be effective for approximating $n-th$ root.

Although primitive approximations such as Babylonian or Bakhshali approximation would be sufficient for engineering or practical purposes, mathematical curiosity and the drive for abstract reasoning compelled mathematicians to invent more sophisticated methods such as Pell's equation, Newton's method, Iterating linear fractions, etc. by forming connections with different branches in mathematics.

Most of the latter methods of approximating square roots were only considered 'mental exercises', as many deemed the calculation of a surd to more than 3 decimal places as an unnecessary task for any real life situations. Yet, some new methods that emerged proved to be more efficient in terms of convergence than primitive methods and thus inadvertently contributed to the development of computational science. Furthermore, through the invention of new methods of approximating square roots, mathematicians could generalize such methods to approximate to the $n-th$ root: a task that could not be accomplished with primitive methods such as the Babylonian and Bakhshali approximation. In addition, new methods such as Newton's and the Bisection method could be combined to form other approximation procedures such as Gould's method. This exemplifies how mathematics is sometimes able to fit in altogether to produce interesting results.

Two major unresolved questions arise from the investigation. First, would comparing more decimal places greatly affect the rate of convergence for $\sqrt{k}$ and $\sqrt[n]{k}$? Also, what exactly is the trend for the convergence for approximating the $n-th$ root? The comparison carried out for the $n-th$ root in this essay was too superficial and carried many assumptions to be accepted as being true.

Overall, from the development of the methods of approximating the square roots of integers to the generalization of the nth root, we can see an example of how a branch of mathematics initially developed from the simple 'need' of people to be able to calculate a certain expression to formulating beautiful methods that allows us to figure out the general method of approximating all $n-th$ root values.

5. Works Cited

Wolfram Alpha Computational Engine . Wolfram Alpha. Ed. Eric Weisstein. Wolfram Alpha. 8 2013 <http://mathworld.wolfram.com/GoldenRatio.html>.

Archibald, Raymond Clare. "Babylonian Mathematics ." Chicago Journals 26.1 (1936): 63-81.

Channabasapa, M.N. "On the square root formula in the Bakhshali manuscript ." On the square root formula in the Bakhshali manuscript 11.2 (1975): 113-126.

Elgin, Dave. "Square Root&Cube Root Algorithms ." Mathematics in School 35.1 (2006): 30-31.

H.W.Gould. "Iterative approximation for finding the Nth root of a number." 33.2 (1959): 61-69.

Lawrence, Dr. Snezana. Math is good for you. 2004/2005. math is good for you. 8 2013 <http://www.mathsisgoodforyou.com/AS/surds.htm>.

Mathematics Stack Exchange. Numerical method for finding the square root - Intel Science and Talent Search Evan O'dorney. 27 10 2012. Mathematics Stack Exchange. 6 2013 <http://math.stackexchange.com/questions/222364/numerical-method-for-finding-the-square-root>.

Thorup, Anders. Notes on Automorphic Functions. Ed. NA. Copenhagen : Københavns Universitets Matematiske Institute, 1995.

6. Bibliography

Dae, Hong Sung. The principle of mathematics . Trans. NA. Vol. 2. Sungji publishing , 2010. 16 vols.

Hofer, Roland. "Characterization of Mobius Transformation." American Mathematical Society journal 128.4 (2000): 1197-1201.

Hong Kong Academy for Gifted Education. "Diophantine Equations ." IMO training 2013 Phase 2 (2013): 4.