Newton's polynomial solver

Christopher J Sangwin
School of Mathematics and Statistics, University of Birmingham,
Birmingham, B15 2TT, UK
Telephone +44 121 414 6197, Fax +44 121 414 3389
Email: C.J.Sangwin@bham.ac.uk http://www.mat.bham.ac.uk/C.J.Sangwin/

February 12, 2002

1 Introduction & historical background

The following short note describes a method of Newton for approximating roots of polynomial equations using slide rules. It first appears in Newton's Waste Book, suggested by [9] around 1665. The method appears then to have been explained to John Collins in a letter of 1672, [7]. Collins was interested in finding the volume of liquid in a partially full barrel as a function of depth[1]. This was an important issue in "gauging" – the calculation of various taxes on liquids. Interestingly slide rules were particularly popular amongst customs officials, for whom they provided an instant volume calculation, by visual inspection, and therefore instantly settled the question of how much tax or duty was due. It is reported in [9] that this method was incorporated by Collins into a general method for solving cubic equations. Collins then wrote *In Answer to Mon^r Leibnitz's Letter about Solving a Cubick æquation by Plaine Geometry.* This *Answer* was passed to Oldenburg who in the following letter, dated June 24 1675, quoted in [2, pg 21], but also in [5], communicated the ideas to Leibnitz.

> Mr. Newton, with the help of logarithms graduated upon scales by placing them parallel at equal distances or with the help of concentric circles graduated in the same way, finds the roots of equations. Three rules suffice for cubics, four for biquadratics. In the arrangement of these rulers, all the respective coefficients lie in the same straight line. From a point of which line, as far removed from the first rule as the graduated scales are from one another, in turn, a straight line is drawn over them, so as to agree with the conditions conforming with the nature of the equations; in one of these rules is given the pure power of the required root.

In [2] the theory is described briefly, and reference is made to [8] where *"Newton's mode of solving equations mechanically is explained more fully and with some restrictions, rendering the process more practical"*. In this short paper we do the same, giving worked examples of both the linear versions as described by Newton, and then Newton's circular version. Since

[1] *"Wherefore since I understand your designe is to get a rule for guageing vessells, this Problem having so bad success for yt end I shall in its stead present you with this following expedient"* [7], pg 229.

none of the quoted sources contain an example to the contrary, it may be that the circular variant of this technique has not before been implemented, owing to the previous difficulty in producing concentric circular logarithmic scales. No record of a commercial polynomial solver using a similar method have so far been found. Prototypes, such as that contained in [6], certainly exist.

Newton, clearly states in [9] that the use of this method is to find "*two or 3 of y^e first figures of y^e desired roote*" in preparation for use of other more accurate numerical techniques. That being the case it's absence from recent histories [3, 11, 12] of numerical techniques is surprising.

It is assumed the reader knows the use of a standard logarithmic slide rule. There are many texts describing its use, however most are now out of print. An exception is [1, Chapter 8].

Acknowledgements

I am greatly indebted to Bruce Williams for querying an error in an early web based explanation of Newton's method, for re-kindling my interest in this topic, and generously providing copies of some references.

2 The principle of Newton's solver

The principle of Newton's polynomial solver is explained with reference to Figure 1. We explain the method with a monic cubic polynomial $p(x)$, satisfying $p(0) = 0$:

$$p(x) := a_1x + a_2x^2 + x^3. \tag{1}$$

As will be seen below it is more convenient when using this method to solve $p(x) = a_0$, rather than the more usual algebraic form $p(x) - a_0 = 0$. Neither the fact that the polynomial is cubic or monic (ie the coefficient of x^3 is 1) is necessary. The method will work with an arbitrary polynomial of the form

$$\sum_{n=1}^{N} a_n x^n.$$

Of course, a polynomial of degree N will utilize N logarithmic rules.

Setting up the apparatus

To set up the apparatus we take three identical logarithmic rules and place them in parallel with equal spacing between them. Working from left to right we refer to these rules as the *primary rule*, *secondary rule* etc. We demarcate an arbitrary *base line* perpendicular to these rules and on the base line attach a *rotating cursor line* at a pivot point, which is as far to the left of the primary rule as the rules are apart.

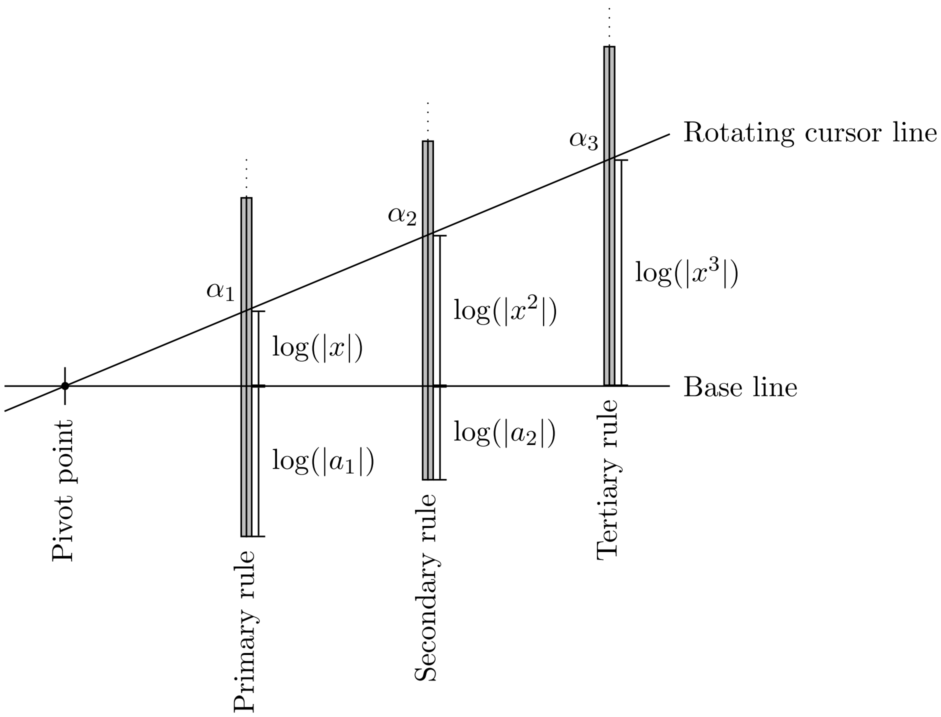


Figure 1: The principle of Newton's polynomial solver

The numerical relationships

We set the rotating cursor line in position and use the primary rule to measure the length between the base line and rotating cursor line. This distance, as measured with the rule, we denote by x. Since the rules are logarithmic the physical distance will be proportional to $\log(|x|)$.

By similar triangles, it is now easy to see that the distance from the base line to the rotating cursor line on the secondary rule will be twice this, so the the physical distance will be proportional to $2\log(|x|) = \log(|x^2|)$.

In general on the distance from the base line to the rotating cursor line on the nth rule will be proportional to $n\log(|x|) = \log(|x^n|)$.

In addition to this we can move these rules in parallel. This uses the usual principle of the linear slide rule to effectively multiply by a constant. Thus if the primary rule is moved so as the reading on the base line is $|a_1|$ the length of rule below the rotating cursor line will be proportional to $\log(|a_1|) + \log(|x|)$. Thus the reading at α_1 will be $|a_1 x|$. Similarly, on the *nth* rule the reading at α_n will be $|a_n x^n|$.

Evaluating polynomials

Let us consider the quantity

$$\alpha_1 + \alpha_2 + \alpha_3$$

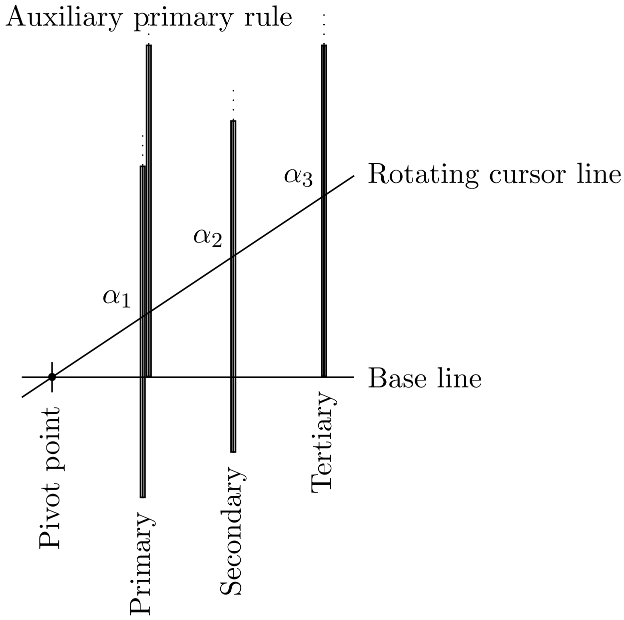


Figure 2: Reading x directly with an auxiliary primary rule.

(In general this corresponds to $\sum_{n=1}^{N} \alpha_n$). Using the above we see that

$$\alpha_1 + \alpha_2 + \alpha_3 = |a_1 x| + |a_2 x^2| + |x^3|.$$

We write $\mathrm{sgn}(w)$ to denote the algebraic sign of w, so that $\mathrm{sgn}(w)|w| = w$. Noting that, since the polynomial is monic so $a_3 = 1$, we have

$$\mathrm{sgn}(a_1)\alpha_1 + \mathrm{sgn}(a_2)\alpha_2 + \mathrm{sgn}(a_3)\alpha_3 = a_1|x| + a_2 x^2 + |x^3|$$

If we assume $x > 0$ then $|x| = x$ so that this reduces to

$$\mathrm{sgn}(a_1)\alpha_1 + \mathrm{sgn}(a_2)\alpha_2 + \mathrm{sgn}(a_3)\alpha_3 = a_1 x + a_2 x^2 + x^3.$$

In summary: if we add up the readings (the α's) on the rotating cursor line, respecting the algebraic signs of the coefficients, we may evaluate the polynomial $p(x)$ for values of $x > 0$.

For values of $x < 0$ we substitute $-x$. Then we have

$$p(-x) = -a_1 x + a_2 x^2 - x^3$$

so that if, in addition to respecting the algebraic signs of the coefficients, we multiply every coefficient of an odd power of x by -1 we evaluate the same polynomial for negative values of x. Examples are below in Section 3.

If $a_1 \neq 1$ the value of $|x|$ may be read directly by use of an auxiliary primary rule, contiguous with the primary rule and aligned with 1 on the baseline itself. See Figure 2. Alternatively we may take the reading from the last scale and extract the cube (or in general nth) root. Since the cubic is monic this will be a pure power of $|x|$.

Solving polynomials

In the previous section we showed that by calculating

$$\operatorname{sgn}(a_1)\alpha_1 + \operatorname{sgn}(a_2)\alpha_2 + \operatorname{sgn}(a_3)\alpha_3$$

and

$$-\operatorname{sgn}(a_1)\alpha_1 + \operatorname{sgn}(a_2)\alpha_2 - \operatorname{sgn}(a_3)\alpha_3$$

for various positions of the rotating cursor line we effectively calculate $p(x)$ and $p(-x)$ respectively for $x > 0$. Thus if we are trying to *solve*

$$p(x) = a_0$$

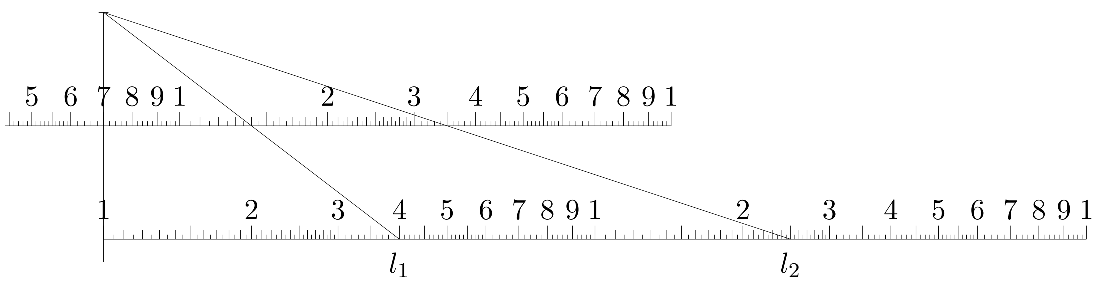
we may do so by finding all positions of the rotating cursor line which creates a sum of the required amount. Note however, that the cases $x > 0$ and $x < 0$ do not require repositioning of the rules themselves. Only that the values read from the rotating cursor be summed differently.

3 Worked examples

3.1 A quadratic with two positive real roots

The following worked example shows the rules set up to evaluate the quadratic

$$p(x) = -7x + x^2.$$



If we examine the two lines, line l_1 gives values

$$\alpha_1 = 14 \text{ and } \alpha_2 = 4$$

Totalling these with the correct signs gives

$$-\alpha_1 + \alpha_2 = -14 + 4 = -10.$$

Similarly line l_2 gives values

$$\alpha_1 = 35 \text{ and } \alpha_2 = 25$$

Totalling these with the correct signs gives

$$-\alpha_1 + \alpha_2 = -35 + 25 = -10.$$

Thus we have found two positions for which $p(x) = -10$. One at $x^2 = 4$, the other at $x^2 = 25$, both with $x > 0$, showing that the solutions to $p(x) = x^2 - 7x = -10$ are $x = 2$ and $x = 5$. Expanding

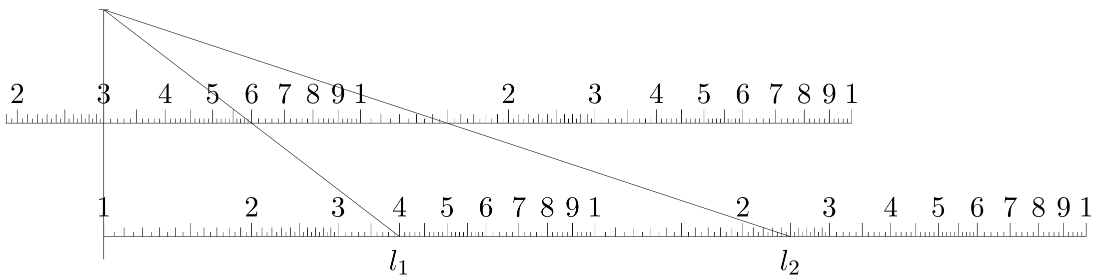
$$(x-2)(x-5) = x^2 - 7x + 10$$

confirms this.

3.2 A quadratic with one positive real root

Next we consider the quadratic

$$p(x) = 3x + x^2.$$



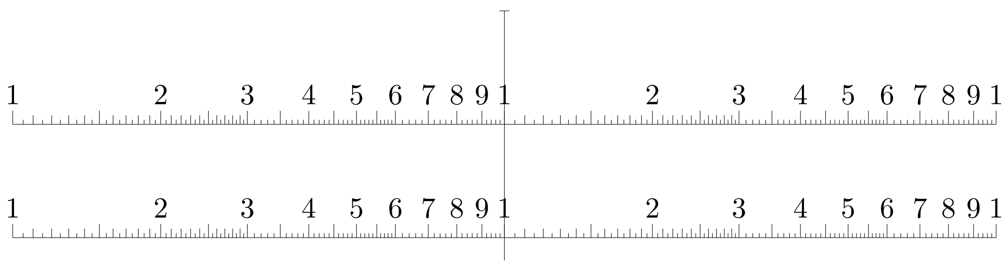
On line l_1 we have $\alpha_1 + \alpha_2 = 6 + 4 = 10$. Similarly on the line l_2 we have $\alpha_1 + \alpha_2 = 15 + 25 = 40$, however, $-\alpha_1 + \alpha_2 = -15 + 25 = 10$. Thus $x = 2$ gives $p(2) = 10$ and $x = -5$ gives $p(-5) = 10$, showing that

$$p(x) - 10 = x^2 + 3x - 10 = (x-2)(x+5).$$

3.3 A quadratic with complex roots

Next, in seeking for solutions to $x^2 - x + 1 = 0$, we consider the quadratic

$$p(x) = -x + x^2.$$

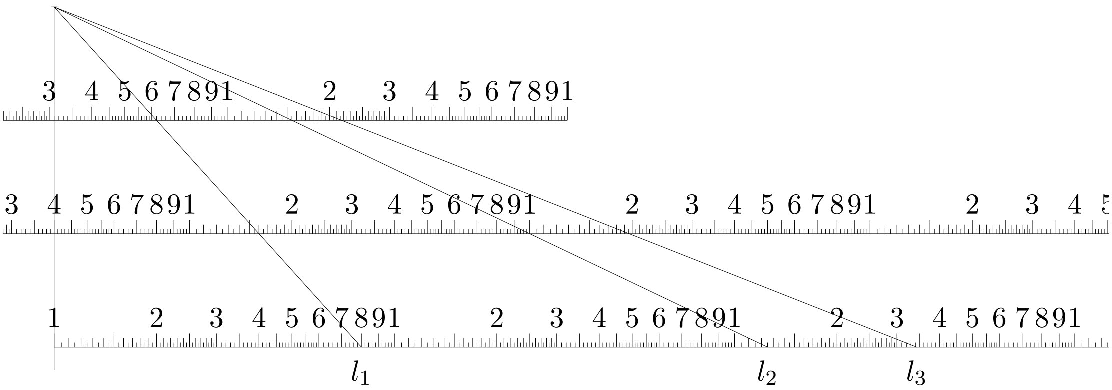


The *target amount* will be values of x so that $p(x) = -1$. If $x < 0$ we sum the readings using signs '$+, +$' so that this will never be negative. There are no solutions to $p(x) = -1$ with $x < 0$. However in the case $x > 0$ we sum the readings using signs '$-, +$'. Thus, does the difference between the bottom and top reading every equal -1? It is far from obvious that this is indeed never the case (since the above quadratic has no real solutions). However in the corresponding case of a cubic with three summands it would be very difficult to conclude the existence of only one real solution.

3.4 Cubic equations

We here consider the cubic polynomial

$$p(x) = -31x - 4x^2 + x^3.$$



Reading from line l_1, from the top down we have readings 62, 16, 8 which, given the signs of the coefficients in the polynomial which are '$-,-,+$' sum as

$$-62 - 16 + 8 = -70.$$

Reading from line l_3, from the top down we have readings[2] 217, 196, 343 which, given the signs of the coefficients in the polynomial which are '$-,-,+$' sum as

$$-217 - 196 + 343 = -70.$$

Returning to line l_2 we have readings $155, 100, 125$. This times we assume $x < 0$ and in addition to the coefficient signs '$-,-,+$' we reverse the signs of the odd coefficients giving '$+,-,-$'. Using this gives

$$155 - 100 - 125 = -70.$$

Taking cube roots of the readings α_3 on each line we see that $x = 2$, $x = 7$ and (remembering in this case $x < 0$) $x = -5$ are solutions of

$$p(x) = -31x - 4x^2 + x^3 = -70$$

which may be confirmed by expanding $(x-2)(x+5)(x-7)$.

4 "Stone's" improvements

In [7], pg 230 Newton remarks in passing at the end that *"As also so to proportion ye rulers BF, CG, DH, &c yt ye line AK may be carried over with parrallel motion"*. This idea is what I term "Stone's improvements" since [2] defers the explanation of them to [8].

[2] To within the accuracy of the scales, I read $220, 200, 342$ giving a sum, $-220 - 200 + 342 = -78$.

In this arrangement, which we will explain with reference to the example of the cubic, the scales used for each power of x differ. Specifically, let the absolute length of the first scale be l. As before this will be the x scale. The length of the second scale will be $l/2$, so that two such scales fit in the same length. This will be the scale for x^2. Such scales on a standard slide rule occupy the B and C positions respectively. The length of the third scale will be $l/3$. Such a scale will be used for x^3, and would be labelled K on a modern slide rule. By placing these rules in parallel at a position with respect to an arbitrary base line, a moving cursor line perpendicular to the rules may be used to evaluate a cubic polynomial.

Firstly, in favour of this method the scales do not need to be equidistant. Secondly the cursor is now parallel the base line – this may facilitate accurate reading.

In opposition to this method we now need a different rule for each power of x. This significantly complicates construction, since there are no repeated parts in the device. Furthermore, scales for each successive power of x are compressed making accurate reading of the scales very difficult for the higher powers.

5 Newton's circular polynomial solver

In the letter of Oldenburg to Leibnitz a scheme for using circular rules is presented. The construction of such an apparatus will require N concentric circular rules. The nth rule will, in one rotation, contain n logarithmic scales. However, the rules need not be equally spaced. A base line is drawn arbitrarily, in our case horizontally, and the reading taken in an anti-clockwise direction.

In Figure 3 we give an example of such an apparatus configured to evaluate the polynomial

$$p(x) = -50x + 35x^2 - 10x^3 + x^4.$$

The 'start' of each logarithmic scale is marked with a dot to facilitate the location of the decimal point. Thus reading across the scales along the base line we have $\alpha_1 = 50$, $\alpha_2 = 35$, $\alpha_3 = 10$, $\alpha_4 = 1$. Summing these, respecting the signs of the coefficients and for the case $x > 0$ in the polynomials we have

$$-50 + 35 - 10 + 1 = -24.$$

The lines l_1, l_2 and l_3 also find values of $x > 0$ for which $p(x) = -24$.

$$\begin{array}{lll} l_1: & -100 + 140 - 80 + 16 & = -24 \\ l_2: & -150 + 315 - 270 + 81 & = -24 \\ l_3: & -200 + 560 - 640 + 256 & = -24 \end{array}$$

Taking the 4th root of the values on the x^4 scale we see that $x = 1$, $x = 2$, $x = 3$ and $x = 4$ all solve

$$p(x) = -24.$$

Hence

$$(x-1)(x-2)(x-3)(x-4) = p(x) + 24 = 24 - 50x + 35x^2 - 10x^3 + x^4.$$

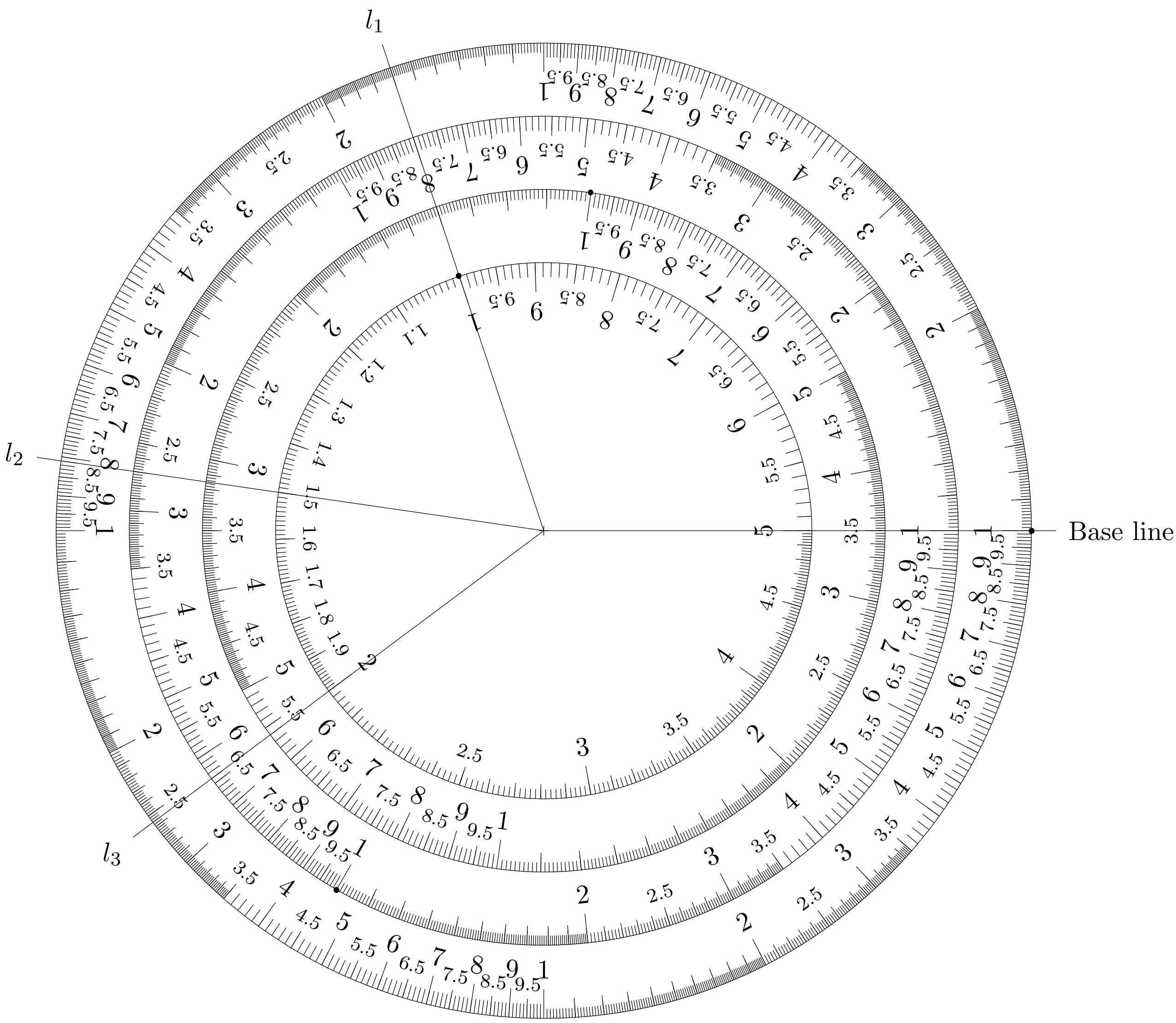


Figure 3: Newton's circular scheme

6 Practical implementation

The apparatus itself would not be hard to build. In the original form described identical scales would be used. A transparent plastic rotating cursor with an opaque ruled line would allow readings to be taken with no more difficulty than on a normal slide rule. Stone's modifications require differing scales – however normal slide rules contain these as C, D and K. The circular arrangement requires different rules and so is perhaps a different matter. Accurately constructing circular scales is not easy, especially without the help of a modern computer typesetting package.

It can readily be confirmed by paper experiments that to use the apparatus is not easy in practice since the implementation requires the simultaneous addition of the coefficients respecting the signs of a_n's and x as appropriate. No doubt a skilled calculator who had carefully planned the calculation could do this. For solving cubic equations a prototype of Stone's modifications has been constructed using three Thornton AD-070 slide rules and is detailed fully[3] in [6]. As suspected it is reported that this is difficult to use.

From paper based experiments, it occurs that the apparatus would be most useful for *evaluating* a polynomial at various values of x, rather than solving the equation $p(x) = a_0$. The solution is mathematically satisfying and clever. One questions whether any application demands the amount of computation involving polynomial evaluation and root finding to within the accuracy permitted by this method. Only such an application would warrant the cost of constructing this apparatus and difficulty in gaining proficiency in its use.

Figures in the current work

This work would not have been possible without the use of a sophisticated mathematical typesetting package. The text is set using LaTeX, the figures themselves use the `PSTricks` and `multido` packages by Timothy Van Zandt, and the `fp` package by Michael Mehlich. LaTeX and associated packages may be obtained from `www.tex.ac.uk`. The author is happy to answer queries about the production of these images, or produce bespoke rules to order.

References

[1] Budd, C. J. and Sangwin, C. J., *Mathematics Galore! Discovery Classes, Workshops and Team Projects in Mathematics and its Applications.* Oxford University Press, 3 May, 2001. ISBN: 0-19-850770-4 (Paperback), ISBN: 0-19-850769-0 (Hardback).

[2] Cajori, F., *A History of the Logarithmic Slide Rule.* J.F. Tapley Co. New York, 1909. Note, this book contains a flawed conclusion as to the priority of the invention of the slide rule. See [4].

[3] Cajori, F., *Historical Note on the Newton-Raphson Method of Approximation*, American Mathematical Monthly, 18(2):29–32 February, 1911.

[3] Also of interest in this article is a method for solving polynomial equations using a standard slide rule. Original sources for this method are British patents numbers 878056/7

[4] Cajori, F., *On the History of Gunter's Scale and the Slide Rule During the Seventeenth Century.* University of California Publications in Mathematics, 1(9):187–209, February 1920.

[5] Horsley S. (ed), *Isaaci Newtoni Opera,* London, VI:520, 1782.

[6] Rath, W., *The Solving of Equations*, Slide Rule Gazette, 2:82–86, 2001. ISSN 1472-0000.

[7] Turn, H. W. (ed), *Correspondence of Isaac Newton — vol 2: 1676–1687* , Cambridge University Press, 229–231, 1960.

[8] Stone, E., *New Mathematical Dictionary.* London, second edition, 1743.

[9] Whiteside, D. T., *The Mathematical Papers of Isaac Newton*, 1, 489–490. Cambridge, 1927.

[10] Wilson, J. *Appendix* to *Mathematical Tracts II*, Robins, B., 437–50, 1761.

[11] Tjalling J. Ypma *Historical Development of the Newton-Raphson Method*, SIAM Review, 37(4):531–551, Dec., 1995.

[12] *Machines for Solving Algebraic Equations,* Mathematical Tables and Other Aids to Computation, 1(9): 337–353, Jan., 1945.