

P1 – Group 7

Line-by-Line Commentary

Discovering the Binomial Theorem

Victor Elgersma & Vera Belde

10 March 2026

Contents

Unpublished Manuscript CUL Add. MS 3958.3, f. 72	2
Integral of the Hyperbola	2
Integral of the Circle	7
Root Extraction: Series from Newton's 'De Analysi'	10
Epistola Prior	12
Epistola Posterior	15

Unpublished Manuscript CUL Add. MS 3958.3, f. 72.¹

In this page from a young Newton's handwritten notes, we find the integrals of $\frac{1}{x}$ and $\sqrt{1-x^2}$ as power series, which Newton 'intuits' by extrapolating the pattern created by a broader class of easier integrals. In his second letter to Leibniz, he explains that this is how he first stumbled upon these series, but notes that polynomial division and the root extraction algorithm provide a firmer foundation.

What follows is not a proof in a sense that a modern mathematician would recognize², but we do think it is pretty convincing.

We first turn to the hyperbola.

Integral of the Hyperbola

If lab is an Hyperbola; cde , ck its Asymptotes, a its vertex, and cag its axis; if $adck$ is a square and he is parallel to ad , and $cd = 1$, and $de = x$, then $be = \frac{1}{1+x}$. If also, $ef = 1$, $eg = 1 + x$, $eh = 1 + 2x + x^2$ etc. (the progression continued is $1 + 3x + 3xx + x^3$, $1 + 4x + 6x^2 + 4x^3 + x^4$, $1 + 5x + 10x^2 + 10x^3 + 5x^4 + x^5$ etc).

This boils down to "draw the curves $y = \frac{1}{x}$, $y = 1$, $y = x'$, $y = x'^2$ and $y = x'^3$ in the $x' - y$ plane. For each of these curves he notes the values at $x' = x + 1$

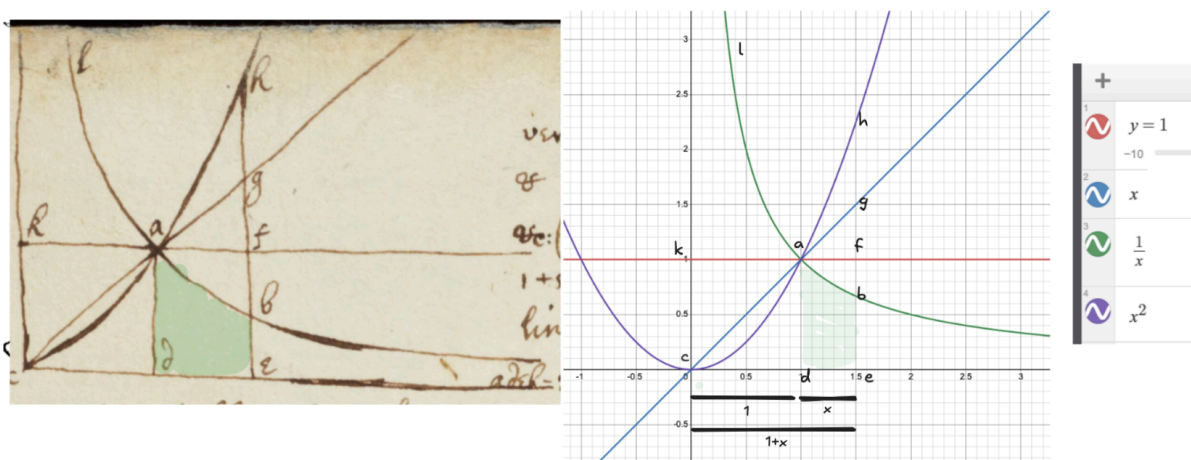


Figure 1: Sketch to show the curves Newton is considering to find the integral of $\frac{1}{x'}$ from 1 to $1+x$, on the right is a facelift done with [Desmos](#)

We use x' and not x because x is reserved for the distance between the points d ($x' = 1$) and e ($x' = 1 + x$). This may seem strange but is necessary for Pascal's triangle coefficients to appear in the integrals that follow.

¹<https://uu.brightspace.com/d2l/le/lessons/44275/topics/498036>, page 191

²For one, Newton manipulates infinite series as if they were finite, without first checking that the series converges

Then, shall the areas of those lines proceed in this progression. $\star = adeb$, $x = adef$, $x + \frac{xx}{2} = adeg$, $adeh = x + \frac{2xx}{2} + \frac{x^3}{3}$, $x + \frac{3xx}{2} + \frac{3x^3}{3} + \frac{x^4}{4}$, $x + \frac{4xx}{2} + \frac{6x^3}{3} + \frac{4x^4}{4} + \frac{x^5}{5}$ etc.

In other words, we calculate the following integrals:

$$\int_1^{1+x} \frac{1}{x'} dx' = \star$$

$$\int_1^{1+x} dx' = x \Big|_1^{1+x} = (1+x) - 1 = 1x$$

$$\int_1^{1+x} x' dx' = \frac{x'^2}{2} \Big|_1^{1+x} = \frac{(1+x)^2 - 1}{2} = 1x + 1\frac{x^2}{2}$$

$$\int_1^{1+x} x'^2 dx' = \frac{x'^3}{3} \Big|_1^{1+x} = \frac{(1+x)^3 - 1}{3} = 1x + 2\frac{x^2}{2} + 1\frac{x^3}{3}$$

$$\int_1^{1+x} x'^3 dx' = \frac{x'^4}{4} \Big|_1^{1+x} = \frac{(1+x)^4 - 1}{4} = 1x + 3\frac{x^2}{2} + 3\frac{x^3}{3} + 1\frac{x^4}{4}$$

$$\int_1^{1+x} x'^4 dx' = \frac{x'^5}{5} \Big|_1^{1+x} = \frac{(1+x)^5 - 1}{5} = 1x + 4\frac{x^2}{2} + 6\frac{x^3}{3} + 4\frac{x^4}{4} + 1\frac{x^5}{5}$$

$$\int_1^{1+x} x'^5 dx' = \frac{x'^6}{6} \Big|_1^{1+x} = \frac{(1+x)^6 - 1}{6} = 1x + 5\frac{x^2}{2} + 10\frac{x^3}{3} + 10\frac{x^4}{4} + 5\frac{x^5}{5} + 1\frac{x^6}{6}$$

We have emphasized the coefficients of Pascal's triangle in **bold**. Newton then draws a table of the coefficients of the integrals he has, which allows him to extrapolate the pattern to establish his best guess for \star :

	$\frac{1}{x}$	1	x	x^2	x^3	x^4	x^5	x^6	x^7
x	?	1.	1.	1.	1.	1.	1.	1.	1.
$\frac{x^2}{2}$?	0.	1.	2.	3.	4.	5.	6.	7.
$\frac{x^3}{3}$?	0.	0.	1.	3.	6.	10.	15.	21.
$\frac{x^4}{4}$?	0.	0.	0.	1.	4.	10.	20.	35.
$\frac{x^5}{5}$?	0.	0.	0.	0.	1.	5.	15.	35.
$\frac{x^6}{6}$?	0.	0.	0.	0.	0.	1.	6.	21.
$\frac{1}{7}x^7$?	0.	0.	0.	0.	0.	0.	1.	7.

Figure 2: Each column represents the integral of a polynomial, where the coefficients are the same as the ones above in **bold**. Rows are increasingly higher terms in the expansions of the integrals. The integral of $\frac{1}{x}$ is the punchline of the 'proof' and we have redacted it at this stage. For the unredacted version see Figure 3.

As in this table. In which the first area is also inserted.

$$\frac{1}{x} \quad 1 \quad x \quad x^2 \quad x^3 \quad x^4 \quad x^5 \quad x^6 \quad x^7$$

x	X	1.	1.	1.	1.	1.	1.	1.	1.
$\frac{xx}{2}$	X	-1.	0.	1.	2.	3.	4.	5.	6.
$\frac{xx^3}{3}$	X	1.	0.	0.	1.	3.	6.	10.	15.
$\frac{xx^4}{4}$	X	-1.	0.	0.	0.	1.	4.	10.	20.
$\frac{xx^5}{5}$	X	1.	0.	0.	0.	0.	1.	5.	15.
$\frac{xx^6}{6}$	X	-1.	0.	0.	0.	0.	0.	1.	6.
$\frac{1}{7}x^7$	X	1.	0.	0.	0.	0.	0.	0.	1.

Figure 3: Same as Figure 2 but with the coefficients for the integral of $\frac{1}{x}$ included. I have added some + and = signs to illustrate Newton’s statement that “The sum of any figure and the figure above it is equal to the figure next to it”

The composition of which table may be deduced from hence; viz: The sum of any figure and the figure above it is equal to the figure following it.

By writing the coefficients in a table and viewing Pascal’s triangle as a square, this allows us to infer a ‘missing’ sequence which in this case coincides with the integral of $1/x$.

Newton doesn’t explicitly state his reasoning for the coefficients of the integral of $\frac{1}{x}$, but we can guess the argument goes something like this: First, the top-left element is probably 1, because all the topmost elements are 1. Second, work out all the subsequent numbers by the rule that “The sum of any figure and the figure above it is equal to the figure following it”. This leads Newton to his punchline:

By which table it may appear that the area of the Hyperbola *adeb* is $x - \frac{xx}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \frac{x^6}{6} + \frac{x^7}{7} - \frac{x^8}{8} + \frac{x^9}{9} - \frac{x^{10}}{10}$ etc.

Newton is guarded in his conclusions. “May appear” is not the same as “has been shown”.³ And indeed, all he has done is provide some numerical evidence that *if* there is a power series expansion for $\frac{1}{x}$, then this particular expansion would respect the pattern.

³This observation was made by Niccolò Guicciardini, Università degli Studi di Milano in his video exposition of this manuscript, available at: <https://cudl.lib.cam.ac.uk/view/MS-ADD-03958/139>

Next, Newton turns his attention to integrating a half-circle between arbitrary points on the x -axis.

Integral of Circle

While Wallis⁴ had expressed π as a power series, Newton generalized his result by finding the integral of a semi-circle between any two points on the x -axis.

Suppose that $adck$ is a Square, abc a circle, age a Parabola, etc. and that $de = x$ and ad is parallel to $fe = 1 = bc$. And that the progression in which the lines fe , be , ge , he , ie , ne , etc. proceeds is $1, \sqrt{1-xx}, 1-xx, \sqrt{1-xx}\sqrt{1-xx}, 1-2xx+x^4, \sqrt{1-2xx+x^4}\sqrt{1-xx}, 1-3xx+3x^4-x^6$, etc.

Like with the hyperbola, Newton begins his study of the area of a circle by drawing an array of curves with the form $(1-x^2)^m$, for both integer and half-integer m .

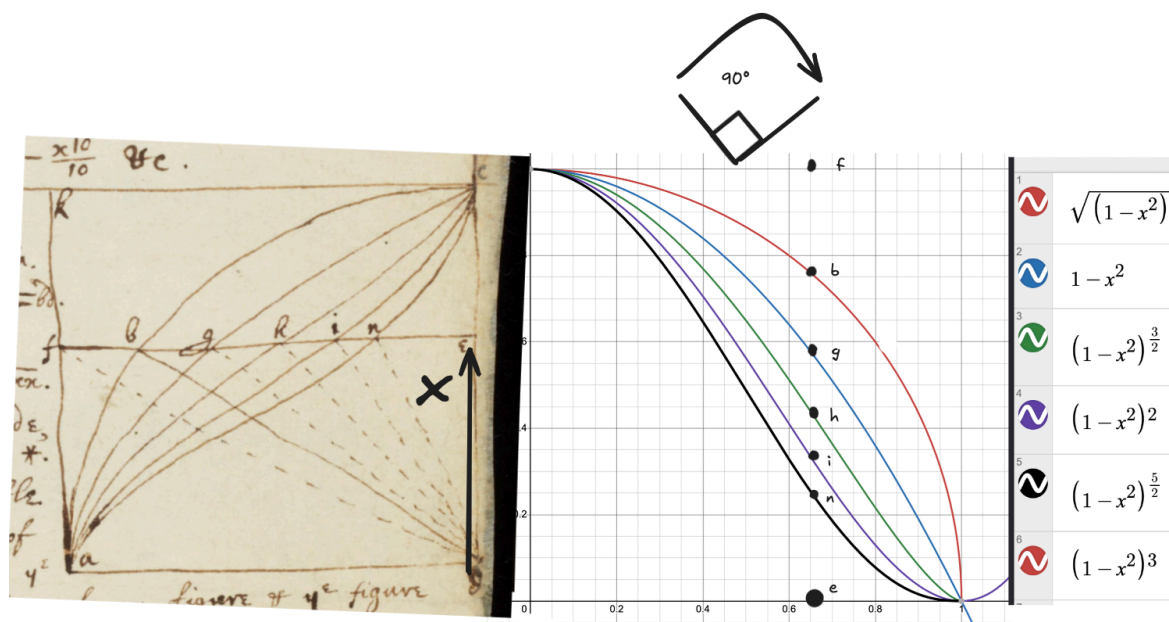


Figure 4: Left: Newton's original sketch. Right: We rotated Newton's sketch 90° clockwise to make sense of it. In modern notation, he is looking for patterns in the integrals of the form

$$\int_0^x (1-x^2)^m dx \text{ for } m = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2} \text{ etc.}$$

Newton then proceeds to integrate the easy ones:

Then will their areas $fade$, $bade$, $gade$, $hade$, $iade$, etc. be in this progression $x, *, x - \frac{xxx}{3}, *, x - \frac{2}{3}x^3 + \frac{1}{5}x^5, *, x - \frac{3x^3}{3} + \frac{3x^5}{5} - \frac{x^7}{7}, *, x - \frac{4x^3}{3} + \frac{6x^5}{5} - \frac{4x^7}{7} + \frac{x^9}{9}$, etc:

In other words, he has just done these integrals:

⁴As claimed in Strogatz, <https://uu.brightspace.com/d2l/le/lessons/44275/topics/498037>

$$\int_0^x (1-x^2)^0 dx = 1x$$

$$\int_0^x (1-x^2)^{\frac{1}{2}} dx = \star$$

$$\int_0^x (1-x^2) dx = 1x + 1\left(-\frac{x^3}{3}\right)$$

$$\int_0^x (1-x^2)^{\frac{3}{2}} dx = \star$$

$$\int_0^x (1-x^2)^2 dx = \int_0^x (1-2x^2+x^4) dx = 1x + 2\left(-\frac{x^3}{3}\right) + 1\frac{x^5}{5}$$

$$\int_0^x (1-x^2)^{\frac{5}{2}} dx = \star$$

$$\int_0^x (1-x^2)^3 dx = \int_0^x (1-3x^2+3x^4-x^6) dx = 1x + 3\left(-\frac{x^3}{3}\right) + 3\frac{x^5}{5} + 1\left(-\frac{x^7}{7}\right)$$

$$\int_0^x (1-x^2)^{\frac{7}{2}} dx = \star$$

$$\int_0^x (1-x^2)^4 dx = \int_0^x (1-4x^2+6x^4-4x^6+x^8) dx = 1x + 4\left(-\frac{x^3}{3}\right) + 6\frac{x^5}{5} + 4\left(-\frac{x^7}{7}\right) + 1\frac{x^9}{9}$$

Again, we have arranged things such that Pascal's triangle coefficients are now in **bold**. Note also the alternating signs between terms.

as in this table following in which the indeterminate terms are inserted.

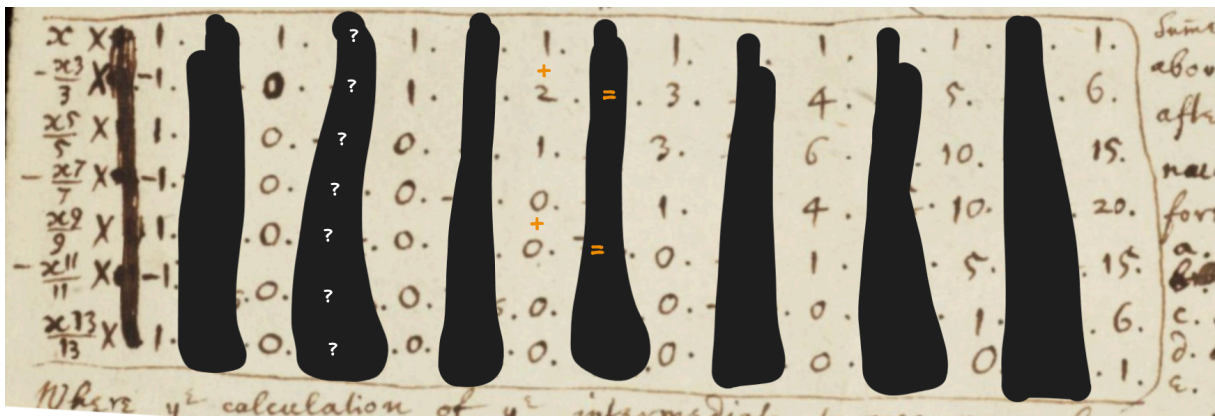


Figure 5: Each column represents the integral of a polynomial of the form $(1-x^2)^m$ for $m = -1, \frac{1}{2}, 0, \frac{1}{2}, 1, \frac{3}{2}$ except we have redacted the values $m = -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}$ etc. The fourth column (redacted) contains the case of the circle $m = \frac{1}{2}$. Newton does not draw the graphs for $m = -1, -\frac{1}{2}$ but he does treat them in this table. Each row represents a term in the series expansion for the aforementioned column.

Like in his treatment of the hyperbola, Newton spots Pascal's triangle:

The property of which table is that the sum of any figure and the figure above it is equal to the figure next after it save one.

Newton then creates the following small finite difference table⁵

Also the numeral progressions are of these forms.

a	a	a	a
b	$a + b$	$2a + b$	$3a + b$
c	$b + c$	$a + 2b + c$	$3a + 3b + c$
d	$c + d$	$b + 2c + d$	$a + 3b + 3c + d$
e	$d + e$	$c + 2d + e$	$b + 3c + 3d + e$

Where the calculation of the intermediate terms may be easily performed.

The image shows a handwritten table with 11 columns and 6 rows. The first column contains powers of x: x^3 , x^2 , x , 1 , 0 , and 0 . The subsequent columns contain numerical values representing differences. The first row of values is: 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1. The second row is: -1, -1/2, 0, 1/2, 1, 3/2, 2, 5/2, 3, 7/2, 4. The third row is: 1, 3/8, 0, -1/8, 0, 3/8, 1, 15/8, 3, 35/8, 6, 63/8, 10, 99/8, 15. The fourth row is: -1, -5/16, 0, 3/48, 0, -1/16, 0, 5/16, 1, 35/16, 4, 105/16, 10, 231/16, 20. The fifth row is: 1, 35/128, 0, -15/384, 0, 3/128, 0, -5/128, 0, 35/128, 1, 315/128, 5, 1155/128, 15. The sixth row is: -1, 63/256, 0, 105/3840, 0, -3/256, 0, 3/256, 0, -7/256, 0, 63/256, 1, 693/256, 6. The seventh row is: 1, 231/1024, 0, -945/46080, 0, 7/1024, 0, -5/1024, 0, 7/1024, 0, -21/1024, 0, 231/1024, 1.

Figure 6: The unredacted table which includes the intermediate terms

We do not see how the calculation of the intermediate terms ‘may be easily performed’ from the finite difference table unless we take into account Newton’s second letter to Leibniz, in which he explains that he obtained a general formula for the binomial coefficients for integer m , and then plugged in $m = \frac{1}{2}$ to get the value for the circle.

The logic goes something like this: First, the coefficient of x in the expansion of $\sqrt{1 - x^2}$ must be 1, since the first row consists entirely of 1s. So $a = 1$. The coefficient of $-x^3$ must be $\frac{1}{2}$ because the second row is an arithmetic series $-1, 0, 1, 2, 3, 4$. So $b = \frac{1}{2}$. Next, we spot the following formula for generating successive terms of Pascal’s triangle. If the second row is m , then the third term is $m \times \frac{m-1}{2}$, the fourth term is $m \times \frac{m-1}{2} \times \frac{m-2}{3}$, the fifth term is $m \times \frac{m-1}{2} \times \frac{m-2}{3} \times \frac{m-3}{4}$ and so on. For a fuller explanation of this please see our section on Newton’s Epistola Posterior.

This is what we would today recognize as the binomial coefficient, i.e the statement that the k^{th} term of the binomial expansion of $(1 - x^2)^m$ is: $\frac{m!}{(m-k)!k!} (-1)^k \frac{x^{2k+1}}{2k+1}$.

⁵We follow Stedall in calling this a ‘finite difference table’; although we do not quite understand how this table leads him to conclude the form of the binomial coefficients. We do see that substituting $a=1$ gives us back the coefficients of Pascal’s triangle.

Root Extraction: A Series from Newton's 'De analysi'⁶

From Newton, 'De analysi', 1669, as published in *Analysis per quantitatum series fluxiones, ac differentias*, 1711, 6-7.

In the following extract, Newton shows another way of obtaining infinite series - by adapting an algorithm that was commonly used at the time for getting the square roots of numbers.

Exempla Radicem Extrahendo.

Si fit $\sqrt{aa+xx} = y$, Radicem sic extraho,

$$\begin{array}{r}
 aa + xx \left(a + \frac{x^2}{2a} - \frac{x^4}{8a^3} + \frac{x^6}{16a^5} - \frac{5x^8}{128a^7} \text{ \&c.} \right. \\
 \frac{aa}{0} + \frac{xx}{xx} + \frac{x^4}{2a^2} \\
 \hline
 0 - \frac{x^4}{4a^2} \\
 - \frac{x^4}{4a^2} - \frac{x^6}{8a^4} + \frac{x^8}{64a^6} \\
 \hline
 0 + \frac{x^6}{8a^4} - \frac{x^8}{64a^6} \\
 + \frac{x^6}{8a^4} + \frac{x^8}{16a^6} - \frac{x^{10}}{64a^8} + \frac{x^{12}}{256a^{10}} \\
 \hline
 0 - \frac{5x^8}{64a^6} + \frac{x^{10}}{64a^8} - \frac{x^{12}}{256a^{10}} \\
 \text{\&c.}
 \end{array}$$

Unde

The first thing to note is that the left bracket has nothing to do with bracketing. Instead, it is supposed to separate the square of our answer from our answer:

square of answer

$$\left(aa + xx \right) \left(a + \frac{x^2}{2a} - \frac{x^4}{8a^3} + \frac{x^6}{16a^5} - \frac{5x^8}{128a^7} \text{ etc.} \right)$$

answer

The algorithm goes as follows: We start with guessing a number whose square will equal a^2 . This is obviously a . Then we take our guess (a), square it (a^2), and take it away from our target answer $a^2 + x^2$. The remainder is x^2 . We take the remainder, divide it by $2a$, and add this value to our old guess (a), to produce our new guess ($a + \frac{x^2}{2a}$). Then repeat the process by squaring our new guess ($a^2 + x^2 + \frac{x^4}{4a^2}$) and taking this away from our target (still $a^2 + x^2$). The remainder ($-\frac{x^4}{4a^2}$) is, like before,

⁶from Newton, 'De analysi', 1669, as published in *Analysis per quantitatum series fluxiones, ac differentias*, 1711, 6-7, available on <https://uu.brightspace.com/d2l/le/lessons/44275/topic>, page 194

divided by $2a$ and added to our old guess $(a + \frac{x^2}{2a})$, creating a new guess $(a + \frac{x^2}{2a} - \frac{x^4}{8a^3})$. This process can continue forever.

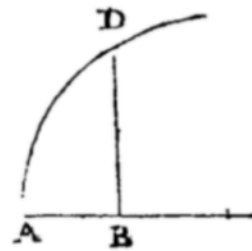
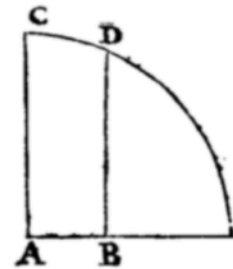
In the next step Newton takes this polynomial expansion of the curve, and integrates it term by term to get the area ABCD.

PER ÆQUATIONES INFINITAS: 7

Unde, pro Æquatione $\sqrt{aa+xx} = y$, nova producitur, viz,

$$y = a + \frac{x^2}{2a} - \frac{x^4}{8a^3} + \frac{x^6}{16a^5} - \frac{5x^8}{128a^7} \text{ \&c. Et (per Reg. 2.) Area quadrata.}$$

$$ABDC \text{ erit} = ax + \frac{x^3}{6a} - \frac{x^5}{40a^3} + \frac{x^7}{112a^5} - \frac{5x^9}{1152a^7} \text{ \&c. Et hæc est Quadratura Hyperbolæ.}$$



“Et hæc est Quadratura Hyperbola” - thus we have squared the hyperbola.

In the following line by line commentary, we will use screenshots from the translation given in *The History of Mathematics: A Source-Based Approach*, Volume 2, chapter 4.2.

Newton's Epistola Prior

This is the first letter Newton sent to Leibniz (via Oldenburg) about the infinite series he found. In this letter Newton is mostly just showing off the series he found and how it can be used.

Fractions are reduced to infinite series by division; and quantities by extraction of the roots, by carrying out those operations in the symbols just as they are commonly carried out in decimal numbers. These are the foundations of these reductions: but extractions of roots are much shortened by this theorem,

$$(P + PQ)^{m/n} = P^{m/n} + \frac{m}{n}AQ + \frac{m-n}{2n}BQ + \frac{m-2n}{3n}CQ + \frac{m-3n}{4n}DQ + \text{etc.}$$

where $P + PQ$ signifies the quantity whose root or even any power, or the root of a power, is to be found; P signifies the first term of Here Newton states what is now known as the Binomial Theorem. However, he uses recursive notation, where A is equal to the first coefficient, here $P^{m/n}$, B is equal to the second coefficient, $\frac{m}{n}AQ = \frac{m}{n}P^{m/n}Q$, and so on. Nowadays, we use the following notation:

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k$$

where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ is known as the binomial coefficient. While in Newton's notation the pattern of the series is very clear and intuitive, the modern notation allows you to directly find the coefficients of a certain term $x^a y^b$ without having to calculate all previous terms. However, Newton's notation allows for non-integer powers, it specifically encourages rational numbers. For the modern notation, you would need to change the binomial coefficient to allow for arbitrary upper index. For an arbitrary number r , you can define:

$$\binom{r}{k} = \frac{r(r-1) \dots (r-k+1)}{k!}$$

and the sum would change from $k = 0$ to r , to $k = 0$ to ∞ . This notation also holds if r is integer, because then the coefficients for $k > r$ would all be zero.

or the root of a power, is to be found; P signifies the first term of that quantity, Q the remaining terms divided by the first, and m/n the numerical index of the power of $P + PQ$, whether that power is integral or (so to speak) fractional, whether positive or negative. For as analysts, instead of aa, aaa , etc., are accustomed to write a^2, a^3 , etc., so instead of $\sqrt{a}, \sqrt{a^3}, \sqrt{c : a^5}$, etc. I write $a^{1/2}, a^{3/2}, a^{5/3}$, and instead of $1/a, 1/aa, 1/a^3$, I write a^{-1}, a^{-2}, a^{-3} . And so for $\frac{aa}{\sqrt{c : (a^3 + bbx)}}$ I write $aa(a^3 + bbx)^{-1/3}$ and for

$$\frac{aab}{\sqrt{c : \{(a^3 + bbx)(a^3 + bbx)\}}}$$

I write $aab(a^3 + bbx)^{-2/3}$: in which last case, if $(a^3 + bbx)^{-2/3}$ is supposed

Here Newton explains his notation. He says he uses the notation of an analyst, the more modern notation, but then for a^2 he continues to use aa , which I found a bit funny, considering he even gave a^2 as an example of notation he would use.

I write $aab(a^3 + bbx)^{-2/3}$: in which last case, if $(a^3 + bbx)^{-2/3}$ is supposed to be $(P + PQ)^{m/n}$ in the Rule, then P will be equal to a^3 , Q to bbx/a^3 , m to -2 , and n to 3 . Finally, for the terms found in the quotient in the

Here Newton explains how this notation fits into his notation of the binomial theorem. In modern notation, you don't need to divide the second term within the brackets by the first term, which does make it a bit easier to use.

m to -2 , and n to 3 . Finally, for the terms found in the quotient in the course of the working I employ A, B, C, D , etc., namely, A for the first term, $P^{m/n}$; B for the second term, m/nAQ ; and so on. For the rest, the use of the rule will appear from the examples.

Here Newton explains how the recursive part of the function works/is used, like I explained in the beginning.

Example 1

$$\sqrt{(c^2 + x^2)} \text{ or } (c^2 + x^2)^{1/2} = c + \frac{x^2}{2c} - \frac{x^4}{8c^3} + \frac{x^6}{16c^5} - \frac{5x^8}{128c^7} + \frac{7x^{10}}{256[c]^9} + \text{etc.}$$

For in this case $P = c^2, Q = x^2/c^2, m = 1, n = 2, A (= P^{m/n} = (cc)^{1/2}) = c, B (= (m/n)AQ) = x^2/2c, C (= \frac{m-n}{2n}BQ) = -\frac{x^4}{8c^3}$; and so on.

Example 2

$$\sqrt[5]{(c^5 + c^4x - x^5)},$$

i.e.

$$(c^5 + c^4x - x^5)^{1/5} = c + \frac{c^4x - x^5}{5c^4} [+] \frac{-2c^8x^2 + 4c^4x^6 - 2x^{10}}{25c^9} + \text{etc.}$$

as will be evident on substituting 1 for m , 5 for n , c^5 for P and $(c^4x - x^5)/c^5$ for Q , in the rule quoted above. Also $-x^5$ can be substituted for P and $(c^4x + c^5)/(-x^5)$ for Q . The result will then be

$$\sqrt[5]{(c^5 + c^4x - x^5)} = -x + \frac{c^4x + c^5}{5x^4} + \frac{2c^8x^2 + 4c^9x + [2]c^{10}}{25x^9} + \text{etc.}$$

Newton then follows the definition with two examples of how his series can be applied. The second example is interesting, for you have multiple options to choose P and Q .

The first method is to be chosen if x is very small, the second if it is very large.

He says when to choose which method but does not explain why. Probably he expects that Leibniz can easily tell what the reason behind this might be. When x is very small, choosing $P = c^5$ allows you to simplify the equation by approximating it only in terms of c , while if x were in the denominator (like when you choose $P = -x^5$), the terms of the series would become bigger and bigger, which would not allow you to accurately approximate the root you are interested in (for Newton here uses it to approximate roots).

As mentioned at the beginning of the line-by-line review of this letter, Newton is in this letter talking about the series he found but not explaining how he found it, or even convinced himself that it works. That happens in the second letter.

Newton's Epistola Posterior

I can hardly tell with what pleasure I have read the letters of those very distinguished men Leibniz and Tschirnhaus. Leibniz's method for obtaining convergent series is certainly very elegant, and it would have sufficiently revealed the genius of its author, even if he had written nothing else. But what he has scattered elsewhere throughout his letter is most worthy of his reputation — it leads us also to hope for very great things from him. The variety of ways by which the same goal is approached has given me the greater pleasure, because three methods of arriving at series of that kind had already become known to me, so that I could scarcely expect a new one to be communicated to us. One of mine I have described before; I now add another, namely, that by which I first chanced on these series — for I chanced on them before I knew the divisions and extractions of roots which I now use. And an explanation of this will serve to lay bare, what Leibniz desires from me, the basis of the theorem set forth near the beginning of the former letter.

This is the introduction of the second letter Newton sent to Leibniz (via Oldenburg) about the infinite series he had found. In this letter, Newton appears much more cordial and explains how he found the series and convinced himself that this series actually works the way he thinks it does. Note that in this letter he does not give any proof, he just explains how he found it.

At the beginning of my mathematical studies, when I had met with the works of our celebrated Wallis, on considering the series by the intercalation of which he himself exhibits the area of the circle and the hyperbola, the fact that, in the series of curves whose common base or axis is x and the ordinates

$$(1 - x^2)^{0/2}, (1 - x^2)^{1/2}, (1 - x^2)^{2/2}, (1 - x^2)^{3/2}, (1 - x^2)^{4/2}, (1 - x^2)^{5/2}, \text{etc},$$

if the areas of every other of them, namely

$$x^3, x - \frac{1}{3}x^3, x - \frac{2}{3}x^3 + \frac{1}{5}x^5, x - \frac{3}{3}x^3 + \frac{3}{5}x^5 - \frac{1}{7}x^7, \text{etc}.$$

could be interpolated, we should have the areas of the intermediate ones, of which the first $(1 - x^2)^{1/2}$ is the circle: in order to interpolate

What Newton is saying here is that he got inspired by the work of Wallis to look at a set of functions, for the area was known if n is even. He wanted to see if he could find a pattern to then also find the area for the functions where n is odd.

ones, of which the first $(1 - x^2)^{1/2}$ is the circle: in order to interpolate these series I noted that in all of them the first term was x and the second terms $\frac{0}{3}x^3, \frac{1}{3}x^3, \frac{2}{3}x^3, \frac{3}{3}x^3$, etc., were in arithmetical progression, and hence that the first two terms of the series to be intercalated ought to be $x - \frac{1}{3}(\frac{1}{2}x^3), x - \frac{1}{3}(\frac{3}{2}x^3), x - \frac{1}{3}(\frac{5}{2}x^3)$, etc. To intercalate

Here Newton is describing how he deduced the first two terms for the series expansion of the area of $(1 - x)^{\frac{n}{2}}, n \in \mathbb{N}$. When n is even, he could expand the area of the function and write down its terms, which allowed him to see a pattern emerge.

lated ought to be $x - \frac{1}{3}(\frac{1}{2}x^3), x - \frac{1}{3}(\frac{3}{2}x^3), x - \frac{1}{3}(\frac{5}{2}x^3)$, etc. To intercalate the rest I began to reflect that the denominators 1, 3, 5, 7, etc., were in arithmetical progression, so that the numerical coefficients of the numerators only were still in need of investigation. But in the alter-

Here he continues his deduction of the pattern for areas beneath functions of the form

$(1 - x)^{\frac{n}{2}}, n \in \mathbb{N}$. He noticed that all the denominators of the coefficient are odd and increasing. Now he only needs to figure out how the numerator evolves.

numerators only were still in need of investigation. But in the alternately given areas these were the figures of powers of the number 11, namely of these, $11^0, 11^1, 11^2, 11^3, 11^4$, that is, first 1; then 1, 1; thirdly, 1, 2, 1; fourthly 1, 3, 3, 1; fifthly 1, 4, 6, 4, 1, etc. And so I began to inquire

This numerical pattern corresponds to what is now called Pascal's triangle (i.e., the binomial coefficients). However, for powers of the number 11, this pattern goes wrong for 11^5 or higher powers due to multi-digit numbers, but not if you carry over the tens digit to the next place value:

- Row 5: 1, 5, 10, 10, 5, 1
- Carry over: 1, (5+1), (0+1), 0, 5, 1 = 161051 = 11^5

And using this, it has been proven that this pattern can be continued, though I personally think that for 11 to powers higher than 4, this is less useful than using the binomial coefficients as presented at the beginning of the first letter. But it is a useful memory aid to check your work.

1, 2, 1; fourthly 1, 3, 3, 1; fifthly 1, 4, 6, 4, 1, etc. And so I began to inquire how the remaining figures in these series could be derived from the first two given figures, and I found that on putting m for the second figure, the rest would be produced by continual multiplication of the terms of this series,

$$\frac{m-0}{1} \times \frac{m-1}{2} \times \frac{m-2}{3} \times \frac{m-3}{4} \times \frac{m-4}{5} \times \text{etc.}$$

Newton here describes the recursive rule that generates the numbers for the numerators. He next gives an example which helps clear it up a bit more.

For example, let $m = 4$, and $4 \times \frac{1}{2}(m-1)$, that is 6 will be the third term, and $6 \times \frac{1}{3}(m-2)$, that is 4 the fourth, and $4 \times \frac{1}{4}(m-3)$, that is 1 the fifth, and $1 \times \frac{1}{5}(m-4)$, that is 0 the sixth, at which term in this case the series stops. Accordingly, I applied this rule for interposing series

Here Newton gives an example to find the binomial coefficients for $(a+b)^4$, where he assumes that you know that the first coefficient is 1 (because for each case he discussed here, the series starts with $(1)x$.

the series stops. Accordingly, I applied this rule for interposing series among series, and since, for the circle, the second term was $\frac{1}{3}(\frac{1}{2}x^3)$, I put $m = \frac{1}{2}$, and the terms arising were

$$\frac{1}{2} \times \frac{\frac{1}{2}-1}{2} \text{ or } -\frac{1}{8}, \quad -\frac{1}{8} \times \frac{\frac{1}{2}-2}{3} \text{ or } +\frac{1}{16}, \quad \frac{1}{16} \times \frac{\frac{1}{2}-3}{4} \text{ or } -\frac{5}{128},$$

and so to infinity. Whence I came to understand that the area of the After Newton saw that his formula worked for the known series, he simply applied the same rules to the series he was interested in and that got him this result. He never gave any proof or reason why he was allowed to do this, he just did it and it worked.

and so to infinity. Whence I came to understand that the area of the circular segment which I wanted was

$$x - \frac{1}{2}x^3 - \frac{1}{8}x^5 - \frac{1}{16}x^7 - \frac{5}{128}x^9 \text{ etc.}$$

And by the same reasoning the areas of the remaining curves, which were to be inserted, were likewise obtained: as also the area of the hyperbola and of the other alternate curves in this series

$$(1 - x^2)^{0/2}, (1 - x^2)^{1/2}, (1 - x^2)^{2/2}, (1 - x^2)^{3/2}, \text{ etc.}$$

And the same theory serves to intercalate other series, and that through intervals of two or more terms when they are absent at the same time. This was my first entry upon these studies, and it had certainly escaped my memory, had I not a few weeks ago cast my eye back on some notes.

But when I had learnt this, I immediately began to consider that the terms

$$(1 - x^2)^{0/2}, (1 - x^2)^{2/2}, (1 - x^2)^{4/2}, (1 - x^2)^{6/2}, \text{ etc.}$$

that is to say,

$$1, 1 - x^2, 1 - 2x^2 + x^4, 1 - 3x^2 + 3x^4 - x^6, \text{ etc.}$$

could be interpolated in the same way as the areas generated by them: and that nothing else was required for this purpose but to omit the denominators 1, 3, 5, 7, etc., which are in the terms expressing the areas; Here Newton remarks how he changed from just looking at an expansion of the area beneath a curve, to the expansion of the curve itself. He noted that he only needed to leave out the denominators and could use the rest of his formula as is.

this means that the coefficients of the terms of the quantity to be intercalated $(1 - x^2)^{\frac{1}{2}}$, or $(1 - x^2)^{\frac{3}{2}}$, or in general $(1 - x^2)^m$, arise by the continued multiplication of the terms of this series

$$m \times \frac{m-1}{2} \times \frac{m-2}{3} \times \frac{m-3}{4} \text{ etc.}$$

so that (for example)

$$(1 - x^2)^{\frac{1}{2}} \text{ was the value of } 1 - \frac{1}{2}x^2 - \frac{1}{8}x^4 - \frac{1}{16}x^6 \text{ etc.,}$$

$$\text{and } (1 - x^2)^{\frac{3}{2}} \text{ was the value of } 1 - \frac{3}{2}x^2 + \frac{3}{8}x^4 + \frac{1}{16}x^6 \text{ etc.,}$$

$$(1 - x^2)^{\frac{1}{3}} \text{ was the value of } 1 - \frac{1}{3}x^2 - \frac{1}{9}x^4 - \frac{5}{81}x^6 \text{ etc.}$$

So now Newton has generalized his formula to not only find the area below certain functions but also find a way to expand the functions themselves.

So then the general reduction of radicals into infinite series by that rule, which I laid down at the beginning of my earlier letter became known to me, and that before I was acquainted with the extraction of roots.

He then generalized that to not only this specific type of function, but any function of the form $(P + PQ)^{m/n}$, which we now know as the binomial theorem.

But once this was known, that other could not long remain hidden from me. For in order to test these processes, I multiplied

$$1 - \frac{1}{2}x^2 - \frac{1}{8}x^4 - \frac{1}{16}x^6, \text{ etc.}$$

into itself; and it became $1 - x^2$, the remaining terms vanishing by the continuation of the series to infinity. And even so $1 - \frac{1}{3}x^2 - \frac{1}{9}x^4 - \frac{5}{81}x^6$, etc. multiplied twice into itself also produced $1 - x^2$.

To check his work, he multiplied the infinite series he found for $(1 - x^2)^{\frac{1}{2}}$ with itself, which should give $1 - x^2$, and it did. Thus he had convinced himself that it worked, but to be more rigorous, he also checked it for $(1 - x^2)^{\frac{1}{3}}$, and here it again also worked.

And as this was not only sure proof of these conclusions so too it guided me to try whether, conversely, these series, which it thus affirmed to be roots of the quantity $1 - x^2$, might not be extracted out of it in an arithmetical manner. And the matter turned out well. This was the form of the working in square roots.

$$\begin{array}{r}
 1 \quad -x^2(1 \quad -\frac{1}{2}x^2 \quad -\frac{1}{8}x^4 \quad -\frac{1}{16}x^6, \text{ etc.} \\
 \hline
 1 \\
 0 \quad -x^2 \\
 \quad -x^2 \quad +\frac{1}{4}x^4 \\
 \hline
 \quad \quad -\frac{1}{4}x^4 \\
 \quad \quad -\frac{1}{4}x^4 \quad +\frac{1}{8}x^6 \quad +\frac{1}{64}x^8 \\
 \hline
 \quad \quad 0 \quad -\frac{1}{8}x^6 \quad -\frac{1}{64}x^8
 \end{array}$$

The method that Newton then uses is explained very well in the line-by-line commentary of *De Analysisi*.

After getting this clear I have quite given up the interpolation of series, and have made use of these operations only, as giving more natural foundations. Nor was there any secret about reduction by division, an easier affair in any case.

He ends his letter stating that he now only uses the last method, the one that was explained in the commentary of *De Analysisi*, and not the binomial theorem itself. For him this had “more natural foundations”, because this was also the way to extract decimals for fractions, so to him it felt natural that it could also be used to extract/approximate a root.