Diophantus

Riley Moriss

October 24, 2025

References are [CO22] for the full text translation as well as an expansive introduction and [HE10] for the introduction and commentary. Note that three whole books of Diophantus were found after Heaths book was written and scholarship has thus progressed greatly.

1 Provenance

Diophantus was an Alexandrian likely living around 250 AD. The authoratative modern compilation of ancient manuscripts, for the first 6 books, from which all other modern translations are derived is that of Tanery (1890's). Since his compilation there have been several other books rediscovered, such that we now have 10 out of a total of 13 (as mentioned by Diophantus himself). The manuscripts available to Tanery only contained 6 of the thriteen and their family tree was nicely summarised by him 1.

We can see that the earliest manuscript is from the 13th century, while most are from the 15th and 16th. Hypatia (Died 415 AD), the daughter of Theon of Alexandria the editor of Euclids works, was perhaps the first commentator and it is thought that most of the manuscripts descend from one edited by her (just as most of Euclids were edited by her father).

In the 20th century 4 new books were discovered, they come from an arabic tradition and seem to be not only commentaries on the text but also substantially rewritten. Contrary to Heaths opinion before they were discovered they were the middle books of the 13 (not the last).

Mostly forgotten during the middle ages, this book was rediscovered by Europe in the 17th century, by people like Fermat (many of his famous notes were written in the margin Bachets edition, including his last theorem, appended to Prop 8. Book II), and inspired much of the then budding number theory. The first published edition of the Greek text that appeared in modern times was by Bachet in 1621, this was the edition owned by Fermat, and also had a facing Latin translation.

Heaths edition was in fact the first English translation of Diophantus.

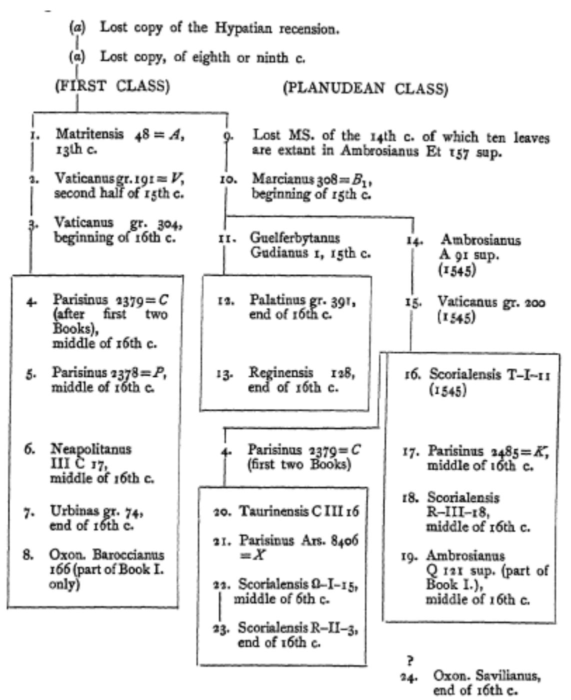


Figure 1:

2 Notation

Nesselmann (I am just cribbing Heath to be clear) classifies the history of algebraic notation into three groups: Rhetorical algebra, syncopated algebra and symbolic algebra. Rhetorical algebra is exemplified by "all Arabian algebraist who are at present known" as well as older Greeks. This was the stage of argumentation where essentially no notation is introduced and the entire proof is carried out in prose, with the exclusion of perhaps numerals (which have essentially always been present). Syncopated algebra is essentially the same except some of the prose is *abbreviated*, that is operations or quantities are replace with a symbol or decoration on other symbols as a shorthand. They are not yet seperated from the prose language of the writer (like for instance x is from representing a number) and will usually denote contractions of a word describing the thing that it denotes. This is the stage to which Diophantus belongs. The final stage is the modern stage where in symbols are completely unrelated to the natural language we use (at least in principle).

Now because Diophantus is in a strange middle ground that we are not really used to in modern times translators are left with three options. One is to regress the writting to the rhetorical stage, this is the approach in [CO22], where the full words are used. The other is to update the notation to modern algebraic standards, this is the approach of [HE10]. One could imagine the third possibility that I do not know of an examples of, that is producing it in abreviated english, Heath gives an example of this on page 51 (apparently such a translation exists in French by Paul van Eecke). As these are the only translations we have and I cannot understand the Greek texts we sort of have to take their word for it. They supplement their translation however with an essay on the actual notation used in the Greek. Again we have to take their word for it.

The first thing to note is that the notation in Diophantus is not systematic. The same thing (eg some operation or say fraction) will be denoted in several different ways, or notation will be replaced with the words etc. There is of course heated debate about which of them is later interpolation etc.

but many seem authentic. Another major limitation is that Diophantus only had one symbol for the unknown, and so could not solve an equation like

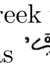
$$ax + b = 0 \implies x = -b/a$$

but instead solved an equation like

$$3x + 7 = 0 \implies x = -7/3$$

(not actually as we will see in the ontology section) the point being that he had to substitute specific numerals for the coefficients. It is stated that he argued in a sufficiently abstract way that this is *merely* a notational limitation, his proof will usually work translated to modern notation in greater generality. Finally worth mentioning is that there is no line spacing. In a modern work one would seperate the steps of their computation onto multiple lines, in Diophantus it is all inline, like prose.

2.1 Variables

In later editions (quasi-modern) the symbol used was $finalsigma$, Nesselman suggests that it was simply the only Greek letter not yet used for a numeral. Heath argues that this is a later innovation and that the symbol used in Diophantus is really a contraction of the first few letters in the greek word "arithmos". The symbol appearing in the manuscript inspected by Heath is drawn by him as .

Regardless of the specific origins of that symbol it was used to denote the unknown. "When the symbol is used in practice the coefficient is expressed by putting the required Greek numeral immediatly after it" 2. Thus we have for instance

$$5x \equiv \varsigma\epsilon$$

Diophantus introduces unique notation for the first six powers of this variable (higher powers just didnt really come up for him it seems). The notation fails to be extensible, the first instances of which apparently took until 1572 by Bombelli. The notation was

$$x^2 \equiv \Delta^Y, x^3 \equiv K^Y, x^4 \equiv \Delta^Y\Delta, x^5 \equiv \Delta K^Y, x^6 \equiv K^Y K$$

"There is no obvious connection between the symbol and the symbol of which it is the square", for this and the reasons above this notation is relegated to the syncopatic stage, as the meaning of these symbols are sought in the actual greek words for power. Again coefficients of these are simply adjoined after the symbol. This same technique was used up until the time of Bombelli, with the Italians calling their variables R, Z, C standing for radix, zensus and cubus (root or unit, square and cube respectively).

One of the great limitations of Diophantus is that he had only one such symbol. This "limitation affects much of (his) work injuriously", because for problems that are most naturally phrased with multiple unknowns he must "assume for one or other some particular number arbitrarily chosen". In some cases II.28,29 he actually performs such a substitution, but tracing it through his argument and once having solved for the first variable then *swaps* in the first solution and replaces his arbitrary substitution for the unknown variable.

2.2 Multiplication

Because Diophantus' variables were all given concrete numerical values he had no need for abstract multiplication notation, the multiplication of the two numbers could be concretely done and multiplication of the variable had its own notation as above.

1	α	alpha	10	ι	iota	100	ρ	rho
2	β	beta	20	κ	kappa	200	σ	sigma
3	γ	gamma	30	λ	lambda	300	τ	tau
4	δ	delta	40	μ	mu	400	υ	upsilon
5	ϵ	epsilon	50	ν	nu	500	ϕ	phi
6	ϛ	vau*	60	ξ	xi	600	χ	chi
7	ζ	zeta	70	o	omicron	700	ψ	psi
8	η	eta	80	π	pi	800	ω	omega
9	θ	theta	90	ϙ	koppa*	900	ϡ	sampi

*vau, koppa, and sampi are obsolete characters

Figure 2:

2.3 Addition

Diophantus uses juxtoposition to denote addition. To be able to add a certain constant number to an expression with variables using this notation he introduces a notation for the unit $\mathring{M}$. Thus he can write

$$x^2 + 3x + 5 \equiv \Delta^Y \alpha\varsigma\gamma\mathring{M}\epsilon$$

2.4 Subtraction

The providence of the subtraction symbol again is an occasion for great scholarly controversy that I will ignore. Some claim it is an upsidedown and truncated Ψ, Heath beleives it is the contraction of the first few letters of the greek word "negation" or "wanting", which was used in their prose discussions of subtraction. The symbol is  however for our convenience we will use $\pitchfork$ in this LaTeX document. Whatever the case because there is no unique sign for addition all the subtractions must be done at once and so all the negative terms are collected on the RHS of a given expression. Thus Diophantus would write

$$x^3 - 5x^2 + 8x - 1 \equiv K^Y \alpha\varsigma\eta \pitchfork \Delta^Y \epsilon\mathring{M}\alpha$$

2.5 Fractions and Division

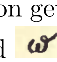
As we will discuss below this is the area in which Diophantus made the most progress and as such there is much to talk about with the surrounding notation.

Sometimes as is the case with multiplication division of numbers can be carried out in a concrete way were the result is a whole number. Otherwise fractions are used to denote the remainder. In some cases the denominator may be omitted entirely. He also often just uses the words "divided by".

When the numerator is 1. For this case there are several simple notations employed, in the form of specific accents

$$\frac{1}{x} = x'' = x' = x^{\wedge} = x^{\times}.$$

This is the notation used for the unknown as well. If more complex expressions in the unknown appear in the denominator then he merely uses natural language.

Other numerators. The first example is 2/3 which for some reason gets its own symbol. It appears four times in Diophantus with a couple of different symbols,  and .

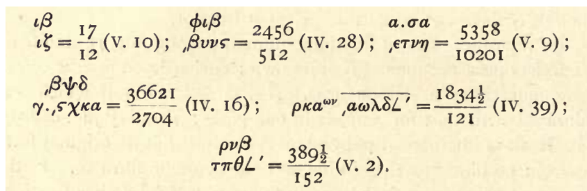
Diophantus also has arbitrary fractions written with the two numbers written one above the other, in his notation however our denominator is written ontop. It is suggested by Heath that in the earliest sources there is no horrozontal line between the two numbers, however in some manuscripts there is. Thus our modern notation is almost the same as the notation appearing in our manuscripts of Diophantus.

In one place in the text fractions are denoted as we now denote exponents, putting the denominator in the power.

The last notation is that of juxtaposing the numerator with the denominator expressed as a fraction with numerator one, that is using the notation in the previous section. For example

$$ab'' = a(1/b)$$

Heath has the nice summary of all the different notations 3.


$\overset{\iota\beta}{\iota\zeta} = \frac{17}{12}$ (v. 10); $\overset{\phi\iota\beta}{,\beta\upsilon\nu\varsigma} = \frac{2456}{512}$ (iv. 28); $\overset{\alpha.\sigma\alpha}{,\epsilon\tau\nu\eta} = \frac{5358}{10201}$ (v. 9);

$\overset{,\beta\psi\delta}{\gamma.,\varsigma\chi\kappa\alpha} = \frac{36621}{2704}$ (iv. 16); $\rho\kappa\alpha^{\omega\nu}\ \overline{,\alpha\omega\lambda\delta\angle'} = \frac{1834\frac{1}{2}}{121}$ (iv. 39);

$\overset{\rho\nu\beta}{\tau\pi\theta\angle'} = \frac{389\frac{1}{2}}{152}$ (v. 2).

Figure 3:

2.6 After Diophantus

Heath goes further and gives a nice survey of the development of our modern algebraic notation that we will paraphrase here. The symbolic stage of algebra was first initiated by the Indians, and appeared in Europe in the 17th century.

$+, -$ first appeared in 1489 due to Johann Widman, although they were not used how we use them now. They reapear in 1518 and are used regularly in Stifels "Arithmetica Integra" (1544).

Robert Recorde (1510 - 1558) in his "The Whetstone of Wit" introduced $=$ but with much longer lines, Vieta (1540 - 1603) used $=$, both notation for equality.

Stifel again used juxtoposition to represent the multiplication of two magnitudes. Harriot (1560-1621) used both juxtaposition and a dot to denote the multiplication of numbers (also introducing $<, >$. Oughtred (1574-1660) used the $\times$ symbol.

$\div$ is found in Rahns algebra (1659).

Descarte in his Geometry (1637) introduced the notation for exponentiation. He also originated the use of x, y, z as the unknowns, where he apparently chose them starting at the end of the alphabet.

3 Ontology

> Diophantus will have in his solutions no numbers whatever except "rational" numbers; and in pursuance of this restriction he excludes not only surds and imaginary quantities, but also negative quantities. Of a negative quantity per se, i.e. without some positive quantity to subtract it from, Diophantus had apparently no conception.

Such solutions to his equations he described as literally "impossible", for instance v.2 where he rejects an equation with a negative solution as absurd.

According to [CO22] "medieval ... algebra operated on a fundamentally different conceptual basis from the algebra initiated by Francois Viete in 1591 and which saw its full expression in Descartes Geometry 1637" and "we include also Diophantus". We now paraphrase some of their notes.

The knowns and unknowns in premodern algebra were always numbers. A "number never means anything other than a definite amount of adefinite objects". Polynomials were derived from the way that numbers were concieved of and represented in arithmetic. We have already noted in our notes on Euclid that in arithmetic the fundamental object was the unit. The unit however was not unique and for the Greeks there "there are indefinitely many units" and a number can be considered as a finite collection of these units. Numbers therefore must be positive and non-zero, as they are collections of units. On the other hand in geometry fractions appear, as well as so called incommensurable. The extent to which these magnitudes can be thought of as numbers per se changes over time. There is one source, that is after Euclid, that attests that at least one author had made the connection between magnitudes, incomensurable lengths and "irrational numbers" (maybe around 100AD).

Diophantus does not seem to deal with irrational numbers, but unlike Euclid he has incorporated rational numbres into his system. Euclid has rational magnitudes and whole numbers, but not rational numbers. We quote Diophantus now

> Since you know that all numbesr are made up of a certain number of units, it is clear that their formation has no limit ... And it is from the addition, subtraction and multiplication of these numbers, and from the ratio which they bear to one another, or, even, each on to its own side, that most arithmetical problems are formed.

Here we can see that rational numbers are included, but Diophantus does invoke geometrical language "one to its own side". Diophantus then introduces ample notation for these numbers further solidifying that they are on a par with the whole numbers.

Despite this conception of numbers already multiplication by negatives was defined in the way we understand it

> A lacking amount multiplied by a lacking amount makes an extant amount; a lacking amount by an extant amount makes a lacking amount.

It is hard for me to see how this makes sense without considering negative quantities as such.

Diophantus states no axioms and makes no reference to Euclids common notions. Is it obvious that they should be assumed, as possed in Euclid, or is Diophantus naive.

4 Remarks

Apparently in the areas of geometry and number theory Euclid and Diophantus can be considered sufficient, they are essentially compendiums of all previous knowledge. In a large part they were so effective that the primary previous sources have been completely lost.

The problem is that the conception of number as a collection of units is really not compatable with rational numbers, thus it is not clear how Diophantus is really thinking about them? One hint we have is that Diophantus dscribes his fractions as

> Every number multiplied by its (reciprocal) makes a unit.

Then it is possible to force the following modern logical structure on Diophantus. Numbers starts as collections of units, we can add collections (essentially disjoint union), itterating this gives us multiplication. In the process of setting up problems and solving them one has to consider inverese operations, this makes sense in the context of subtraction, which at least when subtracting a smaller from a greater can also be thought of intuitively removing some of the units. Geometry gives a precident of two magnitudes having a rational ratio, and if we then define an inverse operation to multiplication one observes certain parallels which ones and there you have the rational number system. It is informed by geometry but formally disjoint from it. It includes rationals in its structure (numbers being not just the units but the resluts of natural operations on the units, originally just addition but then subtraction, multiplication and finally division). To what extent is this true, or evern knowable?

5 Example

Lets consider proposition 8 from book 2. Diophantus states the goal

> To divide a proposed square (number) into two squares.

we now compare the two translations we have access to. [CO22] has it as follows:

> Now, let it be proposed to divide 16 into two squares.
>
> And let the first be assigned to be 1 Power. Therefore, the other will be 16 units lacking 1 Power. Thus, 16 units lacking 1 Power should be equal to a square. I form the square from (a side of) whatever (multitude) of Numbers lacking as many units as there are in the side of the 16 units. Let it be 2 Numbers lacking 4 units. Therefore, the square itself will be 4 Powers, 16 units lacking 16 Numbers. These are equal to 16 units lacking 1 Power.
>
> Let the lacking be added in common and likes from likes. Thus, 5 Powers are equal to 16 Numbers, and the Number becomes 16 5ths.
>
> The one will be 256 25ths, the other 144 25ths, and the two added together make 400 25ths, that is 16 units, and each of them is a square.

Lets go through this (Thanks to Tianqi and Kwan for helping me decipher this): We want to find an $x, y \in \mathbb{Q}$ such that $16 = x^2 + y^2$. Then for the first unknown term (which is a power) we see that it is given by $x^2 = 16 - y^2$, that is 16 lacking one (the coefficient) power (the type). The square that forms the 16 is that given by 4^2, that is it is the area of a square with side length 4, thus "as many units as there are in the side of 16 units" is 4. We now form another square whose sides are "whatever numbers", that is ay lacking the units of the sides of 16, or in other words $ay - 4$. Thus we are considering $(ay-4)^2$. If we let $a = 2$ then this expands to $4y^2 + 16 - 16y$, that is "4 powers, 16 units lacking 16 numbers". If we set $x = 2y - 4$ and we see that $4y^2 + 16 - 16y = 16 - y^2$, "letting the lacking be added in common and likes from likes" we get that $5y^2 = 16y$, "thus 5 powers are equal to 16 numbers". Thus "the number becomes 16 fifths" or $y = 16/5$. We then can compute $x^2 = 16 - (16/5)^2$ and we are done.

One thing to note is that if we chose $a = 1$ then we would have got $x = 0$, however higher values of a are valid. Another thing is that there doesnt seem any principled reason to choose this $x = ay - 4$, it is just sort of a neat trick that its coefficients when squared will cancel off with the other equation to give something of the form $y = r$. Maybe this is principled enough for arithmetic.

It is interesting to think how one might show this in modern language to compare. Here is a proof. We can choose a Pythagorean triple

$$a^2 + b^2 = c^2, \ a, b, c \in \mathbb{N}$$

and then just multiply the right hand side to get it to be 16

$$\implies \left(\frac{4a}{c}\right)^2 + \left(\frac{4b}{c}\right)^2 = a^2\left(\frac{4}{c}\right)^2 + b^2\left(\frac{4}{c}\right)^2 = c^2\left(\frac{4}{c}\right)^2 = 16$$

Diophantus numbers correspond in this proof to the Pythagorean triple $3^2 + 4^2 = 5^2$. We can also see here a little more clearly that the original number needs to be a square so that we can factor the exponent over the summands rationally.

Now recall that Diophantus only has one symbol for the unknown, how did this argument then play out. Heaths translation shows a bit more clearly perhaps than the pure prose above how he is arguing:

Given square number 16.

x^2 one of the required squares. Therefore $16 - x^2$ must be equal to a square.

Take a square of the form[1] $(mx-4)^2$, m being any integer and 4 the number which is the square root of 16, *e.g.* take $(2x-4)^2$, and equate it to $16-x^2$.

Therefore $4x^2 - 16x + 16 = 16 - x^2$,

or $5x^2 = 16x$, and $x = \frac{16}{5}$.

The required squares are therefore $\frac{256}{25}$, $\frac{144}{25}$.

We can see that he labels one of the squares and just refers to the other in prose, "one of the required". Thus he works out equations in two unknowns.

References

[CO22] Jean Christianidis and Jeffrey Oaks. *The Arithmetica of Diophantus: A Complete Translation and Commentary*. Routledge, London, 1 edition, November 2022.

[HE10] Thomas Little Heath and Leonhard Euler. *Diophantus of Alexandria; a study in the history of Greek algebra*. Cambridge : University Press, 1910.