Notes on Diophantus

Kirti Joshi

School of Mathematics, Tata Institute of Fundamental Research, Bombay 400 005, India

Introduction

Diophantus is to Number Theory what Euclid is to Geometry. Both have left an indelible mark on their subjects – a mark which has survived several centuries and even inspired many generations of later day mathematicians. Both will be remembered by posterity, at least, by the subjects which have been named after them. After Euclid, we have the notion of Euclidean geometry, while Diophantine geometry – which is an offspring of the marriage of Arithmetic and Geometry, is the apotheosis of Diophantus. While Euclid's 'Elements' is perhaps the only book, other than the Bible, which was studied, and printed so extensively in the Occidental world, Diophantus' *Arithmetica* remained in the mire of obscurity for several centuries – while Europe plunged into the dark ages, until the great *deus ex machina* of 'Renaissance' cleaned the dust of antiquity from this valuable gem and brought forth its shine and lustre. In these 'notes', we wish to give a perspective of Diophantus' life and his work. It will be apparent, from the very title, that this article hardly makes any claims to completeness – for that would be beyond our competence and, perhaps, also far beyond the scope of this journal. So let there be no doubt about the nature of this article: I will be content to give an account of Diophantus' work and his times as known to historians of our subject. I also want to point out to the reader that my approach to the problems discussed by Diophantus is coloured by my perceptions and tastes in Number Theory, which show a distinct geometrical leaning. Moreover, instead of merely discussing problems and solutions given by Diophantus, we have decided to choose a few sample problems and discuss them from a more modern perspective.

Protohistory

It will be appropriate that we begin by discussing the epoch during which Diophantus lived. The Macedonian king, Alexander the Great, during his conquests, established a city (*circa* BC 335) in Egypt, now known as Alexandria (in arabic *al-Iskandarīyāh*), and made it into an important naval base. After Alexander's death in BC 323, the Alexandrian empire was divided up amongst his generals; the rule of Egypt passed into the hands of his capable general Ptolemy I Soter (not to be confused with the astronomer Ptolemy who lived several hundred years later). When Ptolemy I Soter ('the saviour[1]') came to power, he made Alexandria his capital. Under his patronage, the city flourished and became a great centre of Hellenic and Semitic studies. He founded the great library at Alexandria and also a *museum*[2].

Ptolemy I, apart from being a great politician and a shrewd diplomat, also appears to have been a lover of the arts and the sciences. This tradition of patronage of the arts and sciences was kept up, after his death, by his son (Ptolemy II Philadephus) and later his grandson (Ptolemy III Euergetes). During the Ptolemaic reign and the period which followed it, the epicentre of Hellenic civilization had shifted from Athens to Alexandria, and remained so for the next five hundred years. It was, perhaps, the golden age of Hellenic Sciences.

When the library and the research institution were founded at Alexandria, Ptolemy invited some of the greatest of Hellenic thinkers, philosophers, mathematicians of that time, to come and set up their schools in Alexandria. For instance, Euclid who lived during BC 300, is believed to have lived and written his 'Elements' in Alexandria[3]. In the reign of Ptolemy III, Eratosthenes – who is well-known for the measurement of the circumference of the earth, was the librarian of the library of Alexandria. With the new institution, manned by the most learned of the Hellenes, and with a library which must have been one of the best equipped libraries of the time: for it is said that at one time this library contained no less than 500,000 'books' or manuscripts, erudition flourished in warm sunshine, albeit a bit too warm at times, of Egypt. One must note that, in those times, building a library did not simply mean collecting books – some of the books were written/created in Alexandria. For instance, the Septuagint translation[4] of the Old Testament, from Hebrew into Greek, (the earliest such work) was carried out here; around the same time Homeric texts were being edited here. Euclid founded a school here and his 'Elements' was probably written here – or in any case first expounded here. Thus the Arts and Sciences flourished in Alexandria.

Diophantus of Alexandria

We do not know if Diophantus was of Hellenic origin, but we do know that he lived and wrote in Alexandria. About his personal life and the period during which he lived, very little is known. We are plagued by ambiguities as soon as we attempt to pin down the exact

period during which he lived. Historians of our subject tell us that he must have lived some time between BC 150 and AD 250. A good span of four hundred years! Even these dates have been gleaned in a rather oblique manner. Diophantus in his tract on 'Polygonal numbers', which we will have occasion to deal with later in this article, quotes a definition from Hypsicles, who is also known as the author of the so-called Book XIV of Euclid's elements (on 'Regular solids'). So surely Diophantus lived after Hypsicles – unless he was a contemporary of Hypsicles. Now Hypsicles lived in the second half of the second century BC. So we get the lower date of about BC 150.

Psellus, a minor mathematician-philosopher (*circa* eleventh century AD) in a letter, mentions that Anatolius, later Bishop of Laodicea (about AD 280), dedicated his work on the 'Egyptian method of reckoning' to Diophantus. So Diophantus must have lived in about AD 250. However, according to the habits of the present day mathematicians (who sometimes dedicate their work to other mathematicians who lived about hundred years back), it is possible perhaps that Diophantus lived a hundred and fifty years before this date! Nicomacheus, who lived in the AD 100, wrote a work called *Introduction to Arithmetic*, and he makes no mention of Diophantus or his work *Arithmetica*. This might appear somewhat singular, for if Diophantus had lived before Nicomacheus, then surely he would have been mentioned in his work – even if this work was a rather elementary introduction to the subject. This leads us to believe that Diophantus probably lived after Nicomacheus. It is conceivable that Diophantus lived earlier and Nicomacheus had not heard of him or of his work. In any case, we will not be able to ascertain this unless some fresh evidence comes up. However we think that if the new evidence comes up, it will probably come from some Arabic texts[5]. It is possible that remaining books of *Arithmetica* were forgotten and later lost after the Hypatian recension of the text.

I will now try and place this epoch in a historical perspective for the sake of the reader. Homer, the *Vyāsa* of Greek poetry, lived in the ninth century BC. Herodotus, the ancient historian of Hellenes lived around BC 480. Aristophanes, the father of European theatrical comedy and satire, lived during the fifth century BC; the philosopher-thinker Socrates was his contemporary, the great tragedians Euripides and Sophocles were either his contemporaries or had lived a few years before him. Euclid appears to have lived in the third century BC. Archimedes was killed in the sack of Syracuse, during the second Punic war, in 212 BC. It is probable that Archimedes made his acquaintance with Eratosthenes at Alexandria[6]. The famous geometer, Apollonius, lived in Alexandria – he was born about 25 years after Archimedes. Heron, the astronomer lived in Alexandria, possibly during the same period as Diophantus, and so did the famous geometer Pappus (who lived in the third century AD). Theon (Hypatia's father) lived in Alexandria during the fourth century, and was responsible for writing an improved version of Euclid's *Elements*. It seems that the Nilotic plains of Egypt were endowed with more than just fertile soil. To contrast with and add to this epochal listing, we note that Sanskrit poet Kālidāsa probably lived between the first and seventh century AD, while the astronomer Varāhamihira must have been at the Gupta court in the fourth or the fifth century AD; and so was Āryabhatta; Brahmagupta is presently dated to have lived during sixth or seventh century AD, Bhāskarāchārya provides us the date himself in his work *Siddhānta-siromani*, to be late twelfth century AD.

While the exact period during which Diophantus lived remains unclear, we however do know that he must have died at the age of 84. This follows from an epigram, probably composed by one of his friends or admirers, which gives his age – the epigram says: 'His boyhood lasted for 1/6 of his life, his beard grew after 1/12 more, he married after another 1/7, his son was born five years later, the son lived to half his father's age, and the father died four years after his son.' This leads to a single linear equation in one variable and can be easily solved to get the answer 84.

Attributed to Diophantus are the following three works:

1. *Arithmetica* – originally in thirteen 'books'.
2. A tract on *Polygonal numbers*.
3. A collection of propositions under the title *Porisms*.

Of these three works, *Porisms* appears to have been lost completely; of the other two, only fragments have survived. We do have the first six books of *Arithmetica*, where Diophantus alludes, quite frequently, to the *Porisms*. Some later commentators/translators of *Arithmetica* include a seventh book which appears to be a fragment of the tract on *Polygonal numbers*. It is not clear at what point in time the remaining parts of his work were lost. One of the earliest commentators on the work of Diophantus appear to have been Hypatia, the daughter of Theon of Alexandria. She was murdered by Christian fanatics in AD 415. Her commentary, however, covers only the first six books of *Arithmetica*. After Hypatia's work – which has not survived, (we know of her commentary on Diophantus from some references to it elsewhere) the next commentators on Diophantus were Arabian mathematicians. The lost books of *Arithmetica* may have been lost very early on, because there is no sign that the Arabians possessed these at all. In the *Fihrist*, we are told that Abul Wafa al-Buzjani, who lived in AD 940–987, wrote a commentary (*tafsir*) on his work. There were other Arabian mathematicians who wrote commentaries on

Diophantus' algebra. The *Fakhari*, an algebraical treatise by Abu Bekr Muhammad bin al-Hasan al-Karkhi (*circa* AD 1000), contains a collection of problems in determinate and indeterminate analysis which shows that their author had delved deep into Diophantus, and in fact some of the problems are lifted straight from Diophantus. Whether this is because he realized that these as significant cases of problems under considerations or whether he was borrowing from Diophantus, we may not learn, but even the mention of these problems in this treatise at once indicates the scholarship of its author. The revival of sciences in Europe during renaissance owes a great deal to the sedulous zeal with which Arabs, translated into Arabic, the Greek and Sanskrit texts on astronomy and mathematics[7]. One should also recall that Arabians were also responsible for the transport of Indian notion of zero and the decimal place value notation across the Mediterranean.

The credit of reviving Diophantus during renaissance goes to Regiomontanus. In connection with lectures on Astronomy which he delivered at Padua, he pointed out that 'no one has yet translated from Greek into Latin, the fine thirteen books of Diophantus, in which the very flower of the whole of arithmetic lies hid, the *ars rei et census* which today they call by the Arabic name of algebra'. This was around the year 1462. A good span of a millennium after Hypatia! In correspondence with a certain Bianchini, Regiomontanus expressed his desire to translate from Greek, all the books of Diophantus. Soon he found a manuscript in Venezia (Venice) which contained only the first six books of Diophantus, but the quest failed to yield anything further.

Despite the attention drawn to Diophantus' work by Regiomontanus, it was not until a hundred years later, in the third quarter of the sixteenth century, that Rafael Bombelli found a manuscript on Diophantus in the Vatican library, and conceived the idea of publishing a translation (*circa* AD 1570). Two years later he published a book on algebra and in its preface he mentions his discovery of Diophantus and talks of the attempts he and a certain Antonio Maria Pazzi had made in translating the first six books of Diophantus, before they were called away from these efforts by some more pressing matters. Bombelli did not publish his translation, but extracted all the problems from the first four books and a few from the fifth and embodied them into his Algebra. Though he did not give credit to Diophantus wherever required, he did keep these problems as close to the original as possible. François Viète, who also wrote a treatise on algebra, published in 1593, borrowed heavily from Diophantus; whether he gathered this material by studying Xylander's translation (which we will discuss in a moment) or studied a Greek manuscript in the royal library in Paris is not clear. In any case neither Bombelli nor Viète acted as translators of Diophantus, but rather used Diophantine methods to bring forth their own discoveries (in algebra) to light.

The next writer we have business with is Wilhelm Holzmann, who published under the Hellenized name of Xylander. Xylander, a man of erudition and a Greek scholar, who took to algebra as a hobby, published the first translation of Diophantus (under the title: 'Diophanti Alexandrini Rerum Arithmeticarum Libri sex . . .'). Though he intended to publish the Greek version at some point, it is now clear that he did not publish the Greek text at all and he died in 1576. While Xylander was a Greek scholar, he was no mathematician and his edition is not free from the pitfalls one would expect in this situation, frequently he misrepresents the author, and often carries over the errors from the original manuscript to his text, but this hardly demerits his efforts: for it was a remarkable achievement to produce a reasonable translation from a single manuscript. The task of clarifying these errors and supplying proofs of many assertions, which Diophantus alludes to as being extant in the *Porisms*, was left to Bachet.

In 1621, appeared the famous Bachet edition of Diophantus, along with the Greek text and commentary. This recension was based on a manuscript which Bachet called 'codex Regius' and which has survived and is supposed to be in the National Library at Paris. This edition has played a remarkable role in history of Number Theory: it was in the margin of his copy of this edition of Diophantus that Fermat made the notorious statement, regarding what is now known as his 'Last Theorem'[8]. It was this edition of Diophantus which launched Fermat. After Fermat's death, his son Samuel Fermat, published this famous copy of Diophantus with Fermat's notes and additions. The significance of this edition will become clear later on when we deal with Fermat. Many more editions, recensions appeared later, but none have played as significant a role in the subsequent history of Number Theory as the three discussed above: Xylander opened the way to Diophantus; Bachet, who along with his commentary, paved the way for Fermat; Fermat's copy of Diophantus and his mathematical correspondence, and some of his notorious claims, became a source of inspiration to some of the greatest Number Theorists.

For the sake of completeness we must make note at this point that the first printed edition of Euclid's elements appeared in AD 1482, only a few years after the first printed edition of the Bible, while the Xylander edition appeared almost a hundred years later.

The arithmetic of Diophantus

For long, Number Theory has been known, more for its seductive charms, than by the number of devotees who flock to worship it. The role of geometry on the other hand has been somewhat more dominating. The

foremost amongst devotees of geometry were the Greeks, who converted the Egyptian art of measuring the earth into a profound intellectual exercise. While Hellenic geometry achieved a high degree of sophistication, Number Theory received very little attention after the followers of Pythagoras and Euclid, and most of it at an elementary level.

If one juxtaposes the problems dealt with by the Greek geometers of the time and the arithmetical problems for the sake of comparison, one soon realizes that the geometrical problems like the problem of squaring a circle, duplicating the cube, trisecting an arbitrary angle[9], were of substantial depth. One is however at a loss to find problems of arithmetical nature of similar depth and scope. The arithmetical sections of Euclid, with the exception of the proof of infinitude of primes and the euclidean algorithm for finding the GCD of two numbers, do not go beyond elementary divisibility properties of numbers. At the risk of inviting the critic's wrath, we might add that, there seems to be not much more than mere 'reckoning' that needs be done in dealing with some of the problems which were considered with at the time.

The first problem of arithmetical nature, which dates to this period, and which strikes us as a problem of incomparable subtlety, is now known as the 'Archimedes Cattle problem', and is believed to have been communicated by Archimedes to Eratosthenes (who was at Alexandria) for the attention of mathematicians at Alexandria. In brief, this problem consists of finding eight unknown quantities, viz. the number of bulls and cows respectively, each of four colours. The first part of the problem connects the eight unknowns by seven simple linear equations. The second part adds two quadratic conditions. If we denote the number of bulls of each colour by W, X, Y and Z, and the cows by the corresponding small letters these equations are:

Problem 1. (Archimedes' cattle problem)

The first part:

1. $W = \left(\frac{1}{2}+\frac{1}{3}\right)X + Y,$
2. $X = \left(\frac{1}{4}+\frac{1}{5}\right)Z + Y,$
3. $Z = \left(\frac{1}{6}+\frac{1}{7}\right)W + Y,$
4. $w = \left(\frac{1}{3}+\frac{1}{4}\right)(X + x),$
5. $x = \left(\frac{1}{4}+\frac{1}{5}\right)(Z + z),$
6. $y = \left(\frac{1}{5}+\frac{1}{6}\right)(Y + y),$
7. $z = \left(\frac{1}{6}+\frac{1}{7}\right)(W + w),$

The second part:

1. $W + X = square,$
2. $Y + Z = triangular\ number.$

Recall that a number of the form $n(n+1)/2$, where n is any whole number, is called a triangular number. The first part by itself admits an easy solution, but the second part is what makes the problem interesting. We will skip details as they are of no relevance here, but it suffices to say that in the process of solving this problem, we are naturally led to the problem of solving the equation:

$$u^2 - 4729494t^2 = 1,$$

in integers u, t such that $t \neq 0$.

This equation is the key to the profundity of the problem. Such equations were later considered by Brahmagupta and a complete general solution given by Bhāskarāchārya–Jayadeva; more generally equations of the form

$$u^2 - dt^2 = 1, \tag{1}$$

where d is a non-square positive integer, are popularly known as Pell's equations (Leonard Euler erroneously called them so; and hence we shall refer to these as the Bhāskarāchārya–Brahmagupta equations), and are the first important (and interesting) examples of problems which Number Theory deals with. After having said this, we must add however, that particular cases of this equation – the ones with small values of d, which very often can be solved by inspection, had been noticed before. But in our opinion this particular problem stands apart for its sheer numerical complexity.

While it is certain that neither Archimedes, nor any one at the time could solve the cattle problem (for the smallest solutions for t, u involves numbers of more than forty digits and the smallest value of W is a number which has more than 200,000 digits!), the very reduction of this problem to a Bhāskarāchārya–Brahmagupta type equation is not without interest and indicates familiarity with rudimentary algebraic manipulations (in some form or the other). Diophantus deals with several such problems, though he does not seem to have any general procedure for solving these equations. We will remark on these in greater detail later on.

Just as Euclid's 'Elements' subsumes (and hence in some sense nullifies) the work of his predecessors, so does the work of Diophantus. The genius of both of these men also lay in the impeccable, logical arrangement of the material which was available to them, filling up any gaps in the extant arguments and supplying better proofs; and finally going beyond the existing boundaries and proving new propositions and theorems.

The first sign of an algebraist at work in Diophantus, is the development of notations for the standard operations of algebra. The work, which is largely concerned with problems of indeterminate analysis, has notations for indeterminates (i.e. variables), a symbol for the minus sign and notations for the powers of indeter-

minates and also for their inverses. This at once gives a degree of sophistication, which is not available, for instance, in geometrical arguments. However, we do not find any explicit notations for multiplication, division and addition. Addition of two terms is merely their juxtaposition, and he groups all the terms of the same sign together in his expressions. He also has notations for fractions – without using an explicit notation for division. There is, of course, no notation for zero; and Diophantus, almost always, seeks solutions in positive rational numbers (= positive fractions), dismisses the negative solutions as being unacceptable. He does not stress the importance and subtleties of finding integer solutions (as opposed to rational solutions) to the problems he considers. This point was later stressed by Fermat. In sharp contrast to his algebra, the algebraic formalism which was developed by Āryabhatta and his school, was far more complete and came very close to the modern one. Diophantus' notations were, to say the least, quite cumbersome. Often, he deals with problems where there are actually two or more variables involved, but he uses the same symbol for the different variables. What is remarkable, however, is that despite these, rather trivial notational limitations, by sheer ingenuity and brilliance, he solves the problem at hand. There is, however, a deeper fault of beauty in Diophantus – he solves the problem at hand, but hardly penetrates the matter to a sufficient depth to give insights into the nature of the problem. His concern, it appears to us, is seldom beyond the solution of the problem at hand.

Diophantus' six books which are available are all in the form of arithmetical problems, roughly thirty problems to a book, arranged according to the nature of difficulties and also by the algebraic type of the equations which are encountered while solving these problems. The first book deals entirely with the problem of solving linear equations in one indeterminate (or variable) of the form $ax + b = c$, where a, b, c are some positive integers. Typically, one is given the values of a, b, c and the problem is to find the corresponding value x. To quote a prototypical example, consider the following problem:

Problem 2 (Book I.7). *From the same (required) number to subtract two given numbers so as to make the remainders have to one another a given ratio.*

Diophantus presents the following solution. Let the required number be x, the two numbers to be subtracted be 100, 20 (say) and let the ratio be 3:1. Then $x - 20 = 3(x - 100)$ and so $x = 140$.

As the reader will recognize, there is hardly any difficulty or ingenuity involved in solving such problems. So we shall pass on to other problems without any further comment. There is also no need to mention the discussion of quadratic equations, where, it appears that Diophantus certainly knew the general principle involved – the one involving completing the squares. However, he seems to have admitted only positive roots of the equation. There is only one equation of degree three discussed by him and it has root, which can be detected with relative ease. Whether he had a general procedure to deal with these equations or not is not very clear.

After these cursory remarks on equations involving a single variable, we come to the most important aspect of Diophantus' work. His study of equations involving several variables. But before we do so, we will digress briefly and introduce a few terms which are used nowadays in Algebraic Geometry. We will be dealing with solutions, in integers or rational numbers, of polynomial equations of the form

$$f(X,Y) = \sum_{i,j} a_{i,j} X^i Y^j = 0.$$

Such an equation represents a curve in the (X, Y)-plane and a solution in rational numbers (resp. integers) to this equation will be referred to as a *rational point* (resp. *integer point*) on the curve described by this equation. Sometimes, especially from the point of view of geometry, it is more convenient to consider the above equation in its homogeneous form:

$$F(X,Y,Z) = \sum_{i+j+k=n} a_{i,j} X^i Y^j Z^k.$$

The integer n is called the degree of the equation, the sum being taken over all non-negative triples i, j, k which satisfy $i + j + k = n$. Note that $F(X, Y, 1) = f(X, Y)$ (and $Z^{\text{degree F}} f(X/Z, Y/Z) = F(X, Y, Z)$) and hence any solution (x_0, y_0) of $f(X, Y) = 0$ gives a solution $(x_0, y_0, 1)$ of $F(X, Y, Z) = 0$. Moreover, any solution of (x_0, y_0, z_0) of $F = 0$ with $z_0 \neq 0$ gives a solution $(x_0/z_0, y_0/z_0)$ of $f = 0$. Furthermore, we can pass from the homogeneous form to the non-homogeneous form by the substitution $X \mapsto x$, $Y \mapsto y$, $Z \mapsto 1$, and the other way round by the substitution: $x \mapsto X/Z$, $y \mapsto Y/Z$ and then clearing denominators.

Thus we can use two forms f and F interchangeably. Moreover, we will identify two rational solutions (x_0, y_0, z_0) and (x_1, y_1, z_1) of $F(X, Y, Z) = 0$ if they differ by a nonzero multiple, i.e. if there is non-zero rational number λ such that $(x_1, y_1, z_1) = (\lambda x_0, \lambda y_0, \lambda z_0)$. Note that as F is homogeneous, for any nonzero λ, $F(\lambda X, \lambda Y, \lambda Z) = \lambda^{\text{degree F}} F(X, Y, Z)$, and so that nonzero multiples of a solution also give a solution of the same equation. Thus, it is very convenient to identify these solutions. We will call the coordinates X, Y, Z as *homogeneous coordinates. Since any homogeneous polynomial F has a tautological solution: F*(0, 0, 0) = 0, *we will discard this tautological solution as it has no significance whatsoever.* We will call (0, 0, 0) the trivial solution to $F = 0$ and henceforth, the expression

'solution of $F = 0$' will always mean non-trivial solution, i.e. when at least one of the X, Y, Z is nonzero. We caution the reader that this terminology does not apply to non-homogeneous polynomials.

The real advantage of using the homogeneous form of the equation and homogeneous coordinates comes when we start investigating geometrical properties. For instance, it is not very hard to prove that any line will intersect any other line in exactly one point (unless the two lines coincide) and that any line will intersect any conic in exactly two points, a cubic in three, etc.

The above considerations also apply to equations in several variables and sets of simultaneous equations. These loci will live in 'projective spaces' of appropriate dimensions. We will say that the homogeneous polynomial $F(X, Y, Z)$ cuts out (or defines) a *curve* in the projective plane given by homogeneous coordinates X, Y, Z. If F is linear in X, Y, Z we will say that $F = 0$ defines a *line*; if F is of degree 2 then we will say that F defines a *conic* (or a *quadric curve*); and similarly a *cubic curve* for a cubic polynomial etc. If F has more than three variables, and is linear in the variables, we will say that F defines a *hyperplane*; and is of degree two then we will say that F defines a *quadric* in those variables. This terminology will be in force throughout this article. Note that in questions where rational solutions are demanded, it makes no difference if we look at the equation in its homogeneous form or not. But in certain problems where integer solutions are required to an inhomogeneous equation, the passage to the homogeneous form is not of much use.

Frequently, the question of rational points on $f = 0$ submits to an easy solution if f is homogeneous, while the question of integer points (when f is inhomogeneous) is often more difficult and displays a deeper level of subtlety. Let me give examples. Consider for instance the general equation of a conic:

$$aX^2 + bXY + cY^2 + dX + eY + f = 0,$$

or in its homogeneous form, the equation

$$aX^2 + bXY + cY^2 + dXZ + eYZ + fZ^2 = 0,$$

(we impose the condition that polynomial $aX^2 + bXY + cY^2 + dXZ + eYZ + fZ^2$ is not a product of two linear factors, for otherwise, the conic given by this equation just consists of a union of two lines whose equations correspond to the linear factors). Observe that any line intersects a conic (I shall use the homogeneous coordinates from now on for all the geometrical assertions we make) in two points. Moreover, if one of these two points is rational and the line has rational slope, the second point of intersection is also a rational point. So it follows that if we are given one rational point on a conic, then by drawing lines from this point, with rational slopes, and finding the second point of the intersection, we get an infinite number of rational solutions, each corresponding to a specific value of the slope. Moreover, any two rational solutions span a line with rational slope, hence every rational solution can be obtained in this way – provided we are given one rational solution. Hence we see that any plane conic whose homogeneous equation has rational coefficients, either has an infinity of solutions (all expressible in terms of a single parameter – the slope) or it has no solutions at all. Let me give some examples: The conic $X^2 + Y^2 + Z^2 = 0$ has no rational solutions at all. For instance, the conic $X^2 + Y^2 = Z^2$ has one rational point $(0, -1, 1)$, taking a line of slope $t = a/b$ which passes through this point, we find the familiar parametrization of the circle: $x = 2ab$, $y = a^2 - b^2$, $z = a^2 + b^2$ (or in its non-homogeneous form: $x = (2t/t^2 + 1)$, $y = (t^2 - 1/t^2 + 1)$ and $z = 1$).

Applying this argument to the Bhāskarāchārya–Brahmagupta equation, it follows that the conic described by Bhāskarāchārya–Brahmagupta equation $X^2 - dY^2 = 1$ admits an infinite number of rational solutions (because it admits one solution, viz. $X = 1$, $Y = 0$). However, the question of finding integer points on this curve, is far more difficult – and the solution hardly obvious. Similarly the linear equation (which represents a line):

$$aX + bY = c,$$

where a, b, c are integers, has an infinity of rational solutions, but it may have either an infinity of integer solutions or none at all depending on whether the GCD of a, b divides c or not.

Armed with this simple terminology and a little bit of coordinate geometry, let us get back to Diophantus. Diophantus, almost always, discusses rational solutions to a wide variety of equations. The simple case of a linear equation, which we have alluded to at the end of the previous paragraph, is not mentioned at all. Perhaps, Diophantus considered it too easy – and the assertion about integer solutions follows from Euclid's algorithm of finding the GCD of any two integers.

In Book II we find the following problem:

Problem 3. *To add the same (required) number to two given numbers so as to make each of them squares.*

This leads to a 'double equation' of the form

$$x + a = u^2,\ x + b = v^2.$$

Eliminating x we get $u^2 - v^2 = a - b$ (we assume without the loss of generality that $a > b > 0$), then upon writing this as $(u + v)(u - v) = a - b$ we may introduce another parameter t with condition that $t \neq 0$ and get the trivial identity:

$$u + v = (a - b)/t,\ u - v = t.$$

Thus it follows that $u = \frac{1}{2}((a-b)/t + t)$ and x can be found from the equation $x + a = u^2$. This is more or less Diophantus' solution to the double equation problem given above, though he gives a solution after giving specific values to a, b, t. The underlying principle is the following: by eliminating x we are led to the conic $u^2 - v^2 = (a-b)$, or equivalently in its homogeneous form, to the equation $u^2 - v^2 - (a-b)w^2 = 0$. Now this conic has a visible rational point $(1, -1, 0)$, and hence by what we have said earlier, it has an infinity of rational points, and every one of these rational points provides a solution of our double equation problem.

The method outlined above, at least in its algebraic form, was no doubt familiar to Diophantus, and he seems to use this method of 'double equations' implicitly in several problems. Diophantus, in fact gives a rule for treating the more general pair of equations: $\alpha x + a = u^2$, $\beta x + b = v^2$. His method depends on the identity $\frac{1}{4}(p+q)^2 - \frac{1}{4}(p-q)^2 = pq$. He states the rule thus 'observing the difference (of the equations), seek two numbers (p, q) such that their product equals this difference; then equate either the square of half the difference of the two factors to the lesser of the expressions or the square of half the sum to the greater'. Then one proceeds as earlier.

Another problem of interest is the following:

Problem 4 (Book III.6) *To find three numbers such that their sum is a square and the sum of any pair is also a square.*

Clearly one has to solve the set of equations

$$x + y + z = u^2,\ x + y = v^2,\ y + z = t^2,\ x + z = w^2.$$

By adding the last three equations, and substituting the first, we get the equation:

$$2u^2 = v^2 + w^2 + t^2.$$

Now this quadric equation has several obvious solutions, for instance, we may take $u = \pm 1$, $v = \pm 1$, $w = \pm 1$, $t = 0$. Any rational point of this quadric gives a solution of our original problem. Thus it suffices to produce rational points on this quadric. Now we appeal to a bit of geometry. First of all we write the equation as follows: $u^2 - v^2 = w^2 - u^2 + t^2$. Thus it follows that the plane defined by the linear equation $u = v$ intersects the quadric in the circle $v^2 = w^2 + t^2$, which has an infinity of rational points, given by the formula, $v = a^2 + b^2$, $w = a^2 - b^2$, $t = 2ab$. This gives the solution $(a^2 + b^2, a^2 + b^2, a^2 - b^2, 2ab)$ of the quadric, and this finishes the problem.

Compare this proof with the following ingenius proof Diophantus supplies: 'Let the sum of all the three be $x^2 + 2x + 1 = (x+1)^2$, sum of the first and the second x^2 and hence the third $2x + 1$; let the sum of the second and the third be $(x-1)^2$. Therefore the first is $= 4x$; the second $= x^2 - 4x$. But the first + third = square, that is, $6x + 1$ = square, say 121. Thus $x = 20$.' This proof amounts to observing that the following algebraic identity is valid:

$$2(x+1)^2 - x^2 - (x-1)^2 = 6x + 1,$$

and so Diophantus reduces the problem to choosing an x such that $6x + 1$ is a square. Now square of any integer of the form $6n \pm 1$ is of this type, and hence again there are infinitely many solutions to the problem. This solution is a typical example of Diophantus' algebraic ingenuity. Note our geometrical solution, does not agree with Diophantus' solution, but in fact proves more: it shows that there is a conic lying on the surface $2u^2 = v^2 + w^2 + t^2$ which has a rational point.

After having given the reader a general flavour of problems, and method of solutions given by Diophantus, let me now move on to some of the important problems. First and the foremost amongst these is the solution of the Bhāskarāchārya–Brahmagupta equation:

$$u^2 - dt^2 = 1.$$

If Diophantus had a method of dealing with these equations, then it must have been expounded in the lost books of 'Arithmetica'. In the extant books there is no explicit mention of this problem. It is not very hard to prove that if (u_1, t_1), (u_2, t_2) are two solutions (in integers) of this equation, then so is

$$(u_1u_2 + dt_1t_2,\ u_1t_2 + t_1u_2)$$

(as can be easily checked). In particular, it follows that given one solution of this equation, we can generate infinitely many solutions by applying the above formula. Diophantus seems to have been aware at least of this fact in a somewhat general form – for in Lemma to problem 15 of Book VI, he says: 'Given two numbers, if when some square is multiplied into one of the two numbers and the other number is subtracted from the product, the result is a square, another square larger than the aforesaid square can always be found which has the same property as the aforesaid square.'

Thus he is aware of the fact that if $ay^2 - b = x^2$ has one solution, say (x_0, y_0), then there is always another solution with a larger value of x. This can be proved as follows: take any solution, (x_1, y_1), of $x^2 - 1 = ay^2$, then $(x_0x_1 + ay_0y_1,\ x_0y_1 + y_0x_1)$ is a solution of $ay^2 - b = x^2$ and with a bigger value of x.

In fact, this procedure of generating new solutions from two given ones, for this equation, is no accident. In modern terms, the set of solutions of this equation $x^2 - ay^2 = 1$ forms a group, which is generated by a single element. Thus there is one fundamental solution which gives rise to all the solutions. The fact that the set of solutions is a group is not very difficult to prove – it follows from the formula given in the previous

paragraph. This fact is also embodied in the solution of this problem ('the chakrāvala method') given by Bhāskarāchārya–Brahmagupta–Jayadeva. Fermat, who lived a few centuries after Bhāskarāchārya, appears to have independently rediscovered the general method of solving this problem. As an aside, I must point out a curious fact: both Fermat[10] and Bhāskarāchārya apply their general principle to solve $u^2 - 61t^2 = 1$. Whether this is just a coincidence or indicates an intermediate link, is not clear. Moreover, this is also the first case where the full power of the method of continued fractions (or chakrāvala) needs to be used. For smaller values of d (i.e. $d < 61$), a solution can be given by 'inspection'.

The last book of the available books of *Arithmetica* consists almost entirely of problems connected with right-angled triangles. In appearance these seem very innocuous, but the general theory is, by far, deeper than all the earlier problems. To quote an important example:

Problem 5. *To find three right angled triangles, with rational sides, and equal areas.*

If we fix the area to be some rational number (or even an integer) say n, then we are required to find three triples (x, y, z) such that $x^2 + y^2 = z^2$ and $1/2xy = n$. It is a pleasant exercise in algebra to show that this problem is equivalent to the following problem:

Problem 6 *(Congruent number problem) Finding three rational squares $v - n$, v, $v + n$ which are in arithmetic progression and have a given common difference n.*

The equivalence of these two problems may be seen as follows: suppose for instance we are given $x < y < z$ a solution to the problem of right-angled triangles, we may take $v = (z/2)^2$, and if we are given v, we may take $x = (v + n)^{1/2} - (v - n)^{1/2}$, $y = (v + n)^{1/2} + (v - n)^{1/2}$, $z = 2(v)^{1/2}$.

Diophantus is in fact asking for three such values of v for some (unspecified) value of n. The above problem is nowadays called the congruent number problem. More specifically, we say that n is a congruent number if we can find a rational right-angled triangle with area n. Stated thus, the problem appears to be first mentioned by Bachet in his addendum to Diophantus, but there are also references to it in an anonymous Arabic manuscript dating to AD 970. Note that Diophantus' problem is weaker than the congruent number problem, because Diophantus does not fix the area n.

The formulae given in the earlier paragraph show that any triple x, y, z, such that $x^2 + y^2 = z^2$ and $xy/2 = n$ gives rise to a rational point on the curve $U^2 = V^3 - n^2V$, or in its homogeneous form $U^2W = V^3 - n^2VW^2$. For we can take $U = (v^3 - n^2v)^{1/2}$, $V = v$, $W = 1$. The cubic curve $U^2W = V^3 - n^2VW^2$ is an example of what is called an *elliptic curve*. Thus Diophantus is in fact asking for three rational points on such an elliptic curve for some value of n. Note that it has some obvious rational points, viz. those given by $(0, 0)$, $(0, \pm n)$. Note that this curve, unlike the case of lines or conics, is given by a degree three equation, and in contrast to all the curves dealt with earlier, this curve can have either a finite number of rational points or an infinite number of them. There is no easy way of deciding if the curve will or will not have infinitely many rational points. Moreover, geometry of this curve, which is richer than that of a conic or a line, comes into play in the *arithmetical* question of infinitude of rational points. The richer geometry provides the set of rational points on this curve with a group structure[11] – moreover, the group law in this case is subtler than the one in the case of Bhāskarāchārya–Brahmagupta equation. But in any case if this group is indeed infinite, then one can find more rational solutions from given ones (analogous to the Bhāskarāchārya–Brahmagupta equation). Moreover, it is possible to prove that if $U^2 = V^3 - n^2V$ has infinite number of rational solutions then n is a congruent number – but this takes a bit more effort to prove.

Diophantus' solution is not without interest. His solution can be described as follows: let a, b, c be such that $ab + a^2 + b^2 = c^2$. Now form the right-angled triangles with sides $(c^2 - a^2, 2ac, c^2 + a^2)$, $(c^2 - b^2, 2bc, c^2 + b^2)$ and $((a + b)^2 - c^2, 2c(a + b), c^2(a + b)^2)$. All three have area $abc(a + b)$. Since the conic $ab + a^2 + b^2 = c^2$ has infinitely many rational points, it follows that all the numbers of the form $abc(a + b)$ are congruent numbers.

Fermat also turned to the congruent number problem and his earliest result, is the assertion that 1 is not a congruent number. Fermat claims to have proven this by his 'method of infinite descent'. In fact, this assertion, as was noticed by Fermat is equivalent to the assertion that the equation $X^4 + Y^4 = Z^4$ has no solutions in integers except those satisfying $XY = 0$. Thus the problem of congruent numbers appears to be at the root of Fermat's notorious assertion that $X^n + Y^n = Z^n$ has no solutions in integers if $n \geq 3$ except those satisfying $XYZ = 0$.

Fermat, in fact, went deeper and discovered the arithmetical principle of descent, and he came very close to discovering the fact that the set of rational points on this curve forms a group – Fermat called his method the method of 'chords and tangents'.

In Book II, we find the following problem:

Problem 7 (Book II.8). *To divide any square into sum of two squares.*

This was an entry which must have caught Fermat's attention very early on. For Fermat has several important observations on problems with similar theme. Moreover, it was as a marginal note to this particular

problem, in which Fermat made the notorious claim, now known as 'Fermat's Last Theorem' – which to this day has eluded some of the most original minds and has proved a veritable face-that-launched-a-thousand-ships in the whole of Mathematics. However, here we wish to draw attention to the quadratic problems handled by Diophantus. The general problem of the theory is the following. Suppose we are given a homogeneous quadratic form with integer coefficients:

$$aX^2 + bXY + cY^2.$$

The problem is to find all numbers which can occur as the values of this form. In other words, given an integer m, we are required to find integers x, y such that the equation

$$ax^2 + bxy + cy^2 = m,$$

holds. It is evident that not every form will represent all the integers. For instance, the form $X^2 + Y^2$ cannot possibly take the value –1. The Bhāskarāchārya–Brahmagupta equation for instance represents 1 as the value of the quadratic form $u^2 - dt^2$. Even the above problem of Diophantus can be transcribed into this new terminology: 'can every square be represented as the value of the form $X^2 + Y^2$?'. Diophantus also asserts that any number which is a sum of two squares, is so, in any number of ways (Book II.9) (here, he also means sum of squares of rational numbers). Even more important is the assertion, which he alludes to in Book III.19, where he says 'If there are two whole numbers which are sums of two squares, their product can be resolved into the sum of two squares in two ways.' This follows in fact, from the following easy identity:

$$(a^2 + b^2)(c^2 + d^2) = (ac + bd)^2 + (ad - bc)^2.$$

This assertion clearly indicates that though he might not have possessed a formal proof of this identity, he was aware of its consequences. Similar more general identities also come up later on in Fermat's work. On this particular problem, Fermat had several important observations: Fermat noted that every prime number which is of the form $4n + 1$ is a sum of two squares; if a prime which is a sum of two squares be multiplied to another which is a sum of two squares, then the product is a sum of two squares. Euler proved later that the only numbers which can be represented by the form $X^2 + Y^2$ are those which are products of primes of the form $4n + 1$, and of even powers of primes of the form $4n + 3$. It is easy to see that any number of the form $4n + 3$ is not a sum of two squares. This follows, for instance, from the fact that any square leaves either 0 or 1 upon division by 4. Problem 9 of Book V seems to indicate that Diophantus was certainly aware of this fact.

Diophantus' problem 11 of Book V, indicates that he was aware of the fact that numbers of the form $8n + 7$ cannot be sum of three squares – though this condition is not complete; the complete condition was given by Fermat, who noted that any number of the form $4^m(8n + 7)$ cannot be a sum of three squares. Gauss later proved that with exception of these numbers, every other number is a sum of three squares. We do not know if Diophantus was aware of the fact that every number is a sum of four squares. However, we must note that problems IV.29, IV.30 and V.14 involve representing numbers as a sum of four squares. In these problems, he needs to divide a given number into sum of four squares and does so without stating any explicit condition which has to be satisfied by such a number, while in places where it is required to divide a number into a sum of two or three squares, he does mention the necessary conditions for doing this. From this it appears that he may have been aware of this fact, but more cannot be said.

Relevance of Diophantus to modern Number Theory

Thus the stage was set for Fermat, and starting from these assertions on squares, he proceeded to prove many more. In particular, in a letter to Pascal (dated September 24, 1654) he asserted that he had a proof by his method of infinite descent of the fact that: every prime number of the form $4n + 1$ is a sum of two squares; every prime number of the form $3n + 1$ can be represented by the form $x^2 + 3y^2$; any prime of the form $8n + 1$ or $8n + 3$ can be represented by the form $x^2 + 2y^2$. Of course, Fermat never wrote up his proofs and almost all these assertions, and many new ones were proved by Euler, Legendre and Lagrange. Fermat also made an even more fantastic assertion in this letter to Pascal that every number is a sum of three triangular numbers, four squares, five pentagonal numbers, etc. This last assertion was proved by Cauchy almost two hundred years after Fermat. Fermat's assertions, which were built on Diophantine foundations, led Euler, Legendre, Lagrange deeper into problems of representing numbers by forms. A few years later, Gauss arrived on the scene and breathed new life into the subject. He settled once and for all the question of representing numbers by quadratic forms in two and three variables; and made detailed investigations into the case of forms in several variables. Gauss' study in turn also led to algebraic number theory – a subject which began properly with Kummer's attempt at proving 'Fermat's Last Theorem', and the parallel work of Eisenstein, Dedekind and Kronecker.

In a different direction, the subject of rational points on elliptic curves, which began with Fermat's work on the congruent number problem, and his investigations of the question of rational points on some other elliptic

curves and his discovery of the 'chord and tangent method' for elliptic curves, which had its roots in Diophantus, has been an active area of research for the last three decades. As regards the congruent number problem, as recently as ten years back, J. Tunnell, showed that if a number was a congruent number then it satisfied certain effectively computable conditions, and showed that if we assumed the truth of a very deep conjecture of Birch and Swinnerton-Dyer[12], then, conversely any number which satisfied these conditions was also a congruent number. This was the first serious progress on the congruent number problem – because it gave a remarkably effective and computable criterion to decide if a given number is congruent. The recent work of Gross, Zagier, Kolyvagin and Rubin has now settled the issue completely. The work of these authors together provides a proof of a weak version of the Birch and Swinnerton-Dyer conjecture in the cases, which are of interest in the congruent number problem.

In yet another direction, from Kronecker's time, began the long courtship of Arithmetic and Geometry, and in due course, with the episcopal benediction of André Weil and Alexander Grothendieck (and many others), this courtship ended in a successful marriage. The offspring of this immensely successful wedlock, which has Diophantus as its great-grandfather, is called Diophantine Geometry.

Notes

1 The greek word *soteria* means salvation, and is the root of the word soterology (in English) meaning the doctrine of salvation

2. The word is not used here in the modern sense· the museum *mouseion* (meaning the house of the Muses) in fact was a university, had a library and a research institution attached to it
3 We learn this from Pappus, who says that Apollinius spent a long time with pupils of Euclid at Alexandria
4. Apparently so called because the translation was carried out by seventy two people.
5. While we have not been able to ascertain the details, a translation of a new Arabian manuscript has been published recently in Paris, which we are told, has fresh material on Diophantus.
6. This probably explains the communication of the famous cattle problem to Eratosthenes.
7. The Arabians appear to have become acquainted with the Indian texts around eighth or ninth century AD, or perhaps a century earlier. It also appears that they translated Āryabhatta's name as Ārjabahr or Ārjabhar; from this to Khowārizmi's al-jebr could possibly be a phonetic mutation brought about over a period of one-and-half centuries!
8. Last Theorem asserts that the equation $X^n + Y^n = Z^n$ for $n \geq 3$ has no solutions in integers except those given by $XY = 0$
9. I might add here that all of these problems have been solved in the last hundred and fifty years or so – and all in the negative Lindemann's proof of transcendence of π proves that the circle cannot be squared, and the remaining two follow as a part of Galois's theory of equations.
10. Fermat did not write down his proof and the first completely formal proof was supplied by means of continued fractions, by Lagrange.
11 Recall that group is a non empty set equipped with an operation which lets us produce a new element of the set from any two given elements, and there is a unique identity element and every element has an inverse element – for example the set of all real numbers form a group under addition with 0 as the identity element.
12. For instance, these conjectures predict that any prime number of the form $8n + 5$ or $8n + 7$ is a congruent number.