
The St. John's Review

Volume XLVIII, number one (2004)

Editor

Pamela Kraus

Editorial Board

Eva T. H. Brann

Frank Hunt

Joe Sachs

John Van Doren

Robert B. Williamson

Elliott Zuckerman

Subscriptions and Editorial Assistant

Audra Price

The *St. John's Review* is published by the Office of the Dean, St. John's College, Annapolis: Christopher B. Nelson, President; Harvey Flaumenhaft, Dean. For those not on the distribution list, subscriptions are \$10.00 for one year. Unsolicited essays, reviews, and reasoned letters are welcome. Address correspondence to the *Review*, St. John's College, P.O. Box 2800, Annapolis, MD 21404-2800. Back issues are available, at \$5.00 per issue, from the St. John's College Bookstore.

©2004 St. John's College. All rights reserved; reproduction in whole or in part without permission is prohibited.

ISSN 0277-4720

Desktop Publishing and Printing

The St. John's Public Relations Office and the St. John's College Print Shop

Contents

Classical Mathematics and Its Transformation

St. John's College, Annapolis

June 5-6, 2004

Foreword.....5
Harvey Flaumenhaft

Why We Won't Let You Speak Of the Square
Root of Two.....7
Harvey Flaumenhaft

The Husserlian Context of Klein's
Mathematical Work.....43
Burt C. Hopkins

Words, Diagrams, and Symbols: Greek and
Modern Mathematics or "On the Need to Rewrite
the History Of Greek Mathematics" Revisited.....71
Sabetai Unguru

A Note on the Opposite Sections and Conjugate
Sections In Apollonius of Perga's *Conica*.....91
Michael N. Fried

Viète on the Solution of Equations and the
Construction Of Problems.....115
Richard Ferrier

 Foreword
Harvey Flaumenhaft

On the weekend of June 5-6, 2004, the Annapolis campus of St. John's College hosted a conference on the topic of classical mathematics and its transformation. This is a pivotal topic in the program of instruction here at St. John's, where classes consist in the discussion of great books ancient and modern, and where approximately half of the entirely prescribed curriculum may be classified as either mathematics or heavily mathematical natural science.

The topic was made truly pivotal at St. John's by Jacob Klein, who joined the faculty in the second year of the New Program (1938-39) and became Dean eleven years later. Just a few years before arriving at St. John's, Klein had published *Die Griechische Logistik und die Entstehung der Algebra* (later translated by another Dean—Eva Brann—as *Greek Mathematical Thought and the Origin of Algebra*). In a letter written just a few weeks after arriving at the College, Klein said, "It is almost unbelievable to me that all the things that occupied me for years, that is, the whole theme of my work, are realized here. The people don't do quite right, very much is superficial and they are not quite right about certain fundamentals. But it...is clear that I am in the right spot." Klein's presence, and his book because of his presence, had a deepening and correcting effect on the life of this college, and through this college, on intellectual life far beyond our halls. The effects of his work pervade the conference papers, which are printed as this issue of *The St. John's Review*.

The conference took place through the generosity of the Andrew W. Mellon Foundation, as the culminating event under a large grant supporting faculty study groups led by me, over several years, on the topic of the conference. The

Harvey Flaumenhaft is Dean at the Annapolis Campus of St. John's College.

grant also supported groups on the same topic on the Santa Fe campus, with participation by faculty members from Thomas Aquinas College. Grateful acknowledgment is due to Sus3an Borden, the college's Manager of Foundation Relations in Annapolis, who did a great deal of work, under very severe time constraints, to set up the conference.



Why We Won't Let You Speak of the Square Root of Two

Harvey Flaumenhaft

Part One

Progress and Preservation in Science

We often take for granted the terms, the methods, and the premises that prevail in our own time and place. We take for granted, as the starting-points for our own thinking, the outcomes of a process of thinking by our predecessors.

What happens is something like this: Questions are asked, and answers are given. These answers in turn provoke new questions, with their own answers. The new questions are built from the answers that were given to the old questions, but the old questions are now no longer asked. Foundations get covered over by what is built upon them.

Progress can thus lead to a kind of forgetfulness, making us less thoughtful in some ways than the people whom we go beyond. We can become more thoughtful, though, by attending to the thinking that is out of sight but still at work in the achievements it has generated. To be thoughtful human beings—to be thoughtful about what it is that makes us

Harvey Flaumenhaft is Dean at the Annapolis campus of St. John's College. This paper was delivered in the form of a Friday-night formal lecture at St. John's College in Annapolis on 27 August 1999.

The first part of it is adapted from the "Series Editor's Preface" in the volumes of *Masterworks of Discovery: Guided Studies of Great Texts in Science* (New Brunswick, NJ: Rutgers University Press, 1993-1997).

The last part is adapted from the "Introductory Note on Apollonius" in Apollonius of Perga, *Conics: Books I-III*, new revised edition, Dana Densmore, ed. (Santa Fe, NM: Green Lion Press, 1998).

The three parts placed between them are adapted from the manuscript *How Much and How Many: The Euclidean Foundation for Comparisons of Size in Classical Geometry* (forthcoming from Green Lion Press), which is part of the larger manuscript *Insights and Manipulations: Classical Geometry and Its Transformation—A Guidebook: Volume I, Starting up with Apollonius; Volume II, From Apollonius to Descartes* (also forthcoming).

human—we need to read the record of the thinking that has shaped the world around us, and still shapes our minds.

Scientific thinking is a fundamental part of that record, but it is a part that is read even less than the rest, largely. That is often held to be because of the opinion that books in science, unlike those in the humanities, simply become outdated: in science, the past is held to be *passé*.

Now science is indeed progressive, and progress is a good thing; but so is preservation. Progress even requires preservation: unless there is keeping, our getting is little but losing—and keeping takes plenty of work.

Precisely if science is essentially progressive, we can truly understand it only by seeing its progress *as* progress. This means that our minds must move through its progressive stages. We ourselves must think through the process of thought that has given us what we would otherwise thoughtlessly accept as given. By refusing to be the passive recipients of ready-made presuppositions and approaches, we can avoid becoming their prisoners. Only by actively taking part in discovery—only by engaging in re-discovery ourselves—can we avoid both blind reaction against the scientific enterprise and blind submission to it.

We and our world are products of a process of thinking, and truly thoughtful thinking is peculiar: it cannot simply outgrow the thinking it grows out of. When we utter deceptively simple phrases that in fact are the outcome of a complex development of thought—phrases like “the square root of two”—we may work wonders as we use them in building vast and intricate structures from the labors of millions of people, but we do not truly know what we are doing unless we at some time ask the questions which the words employed so casually now were once an attempt to answer.

The education of a human being requires learning about the process by which the human race gets its education, and there is no better way to do that than to read the writings of those master-students who have been the master-teachers.

Part Two

Manyness: The Classical Notion of Number

The first great teachers of the West, and subsequently the rest of the world, were the classical teachers who wrote in Greek some two-and-a-half millennia ago. In their language the word for things that are learned or learnable is *mathêmata*, and the art that deals with the *mathêmata* is *mathêmatikê*—the English for which is “mathematics.”

In mathematics, the first and fundamental classic work is the *Elements* of Euclid. From its Seventh Book we learn that a number is a number *of things of some kind*. When there are more than one of something, their number is the “multitude” of them. It tells how many of them there are. Suppose, for example, that in a field there are seven cows, four goats, and one dog. If we count cows, then the cow is our unit, the cow being that according to which any being in the field will count as one item; and there is a multitude of seven such units in the field. If we count animals, however, then the animal is our unit; and now there is a multitude of twelve units. Of cows, there are seven; but of animals, there are twelve.

Of dogs as dogs, there is no reason to make a count (as distinguished from including them in the count of animals). There is no reason to count dogs, for the field does not contain dogs. It contains only one dog, and one dog alone does not constitute any multitude of dogs. If what we were thinking was “There is *only* one dog in the field,” we might say “There is one dog in the field”; but if we were not thinking of that single dog in relation to some possible multitude of dogs, we would simply say “There is *a* dog in the field.” Of pigs, there is no reason to make a count, for the field does not contain a single pig, let alone a multitude of pigs.

If there had earlier been six horses in the field, which were then lent out, the field might now be said to lack six horses; but while six is the multitude of horses that are lacking *from* the field, there is not any multitude of horses *in* the field.

If we chop up one of the cows, arranging to divide its remains into four equal pieces, and we take three of them, then we have not taken any multitude of cows. We have taken merely three pieces that are each a fourth part in the equal division of what is a heap of the makings for beef stew.

And so when we have several items, each of which counts as one of what we are counting, then what we obtain by counting is a number. The numbers, in order, are these: two units, three units, four units, and so on. The numbers, we might even say, are these: a duo, a trio, a quartet, and so on. To us, nowadays, that seems strange: a lot seems to be lacking.

“One” does not name a multitude of units: although a unit can be combined with, or be compared with, any number of units, a unit is not itself a number—so “one” does not name a number. “Zero” also does not name a number, for there is no multitude of units when there are no units for a multitude to be composed of. Neither does “negative-six” name any multitude, although “six” does; “six” names a number, but there is no number named “negative-six”—and hence there is no number named “positive-six” either.

Although “three” names a number of units, and so does “four,” “three-fourths” does not name a number of units—for if the unit is broken up, then, being no longer unitary, it ceases to be a unit. The Latin word for breaking is “fraction.” A fraction is not a number. (Of course we can, if we wish, take a *new* unit—say, a fourth part of a cow’s equally divided carcass—and then take three of these new units.)

As for “the square root of two,” it is not a number; indeed, as we shall soon see, it is not even a fraction. And neither is “pi” a number.

To say it again, the numbers are: two units, three units, four units, and so on—taking as many units as we please. That is what Euclid tells us; and, after telling us, Euclid goes on to speak of the relations among different numbers.

We learn, for example, that because fifteen “measures” sixty—that is to say, because a multitude of fifteen units taken

four *times* is equal to a multitude of sixty units—fifteen is called “part” of sixty, and sixty is called a “multiple” of fifteen.

But fifteen is not “part” of forty, because you cannot get a multitude of forty units by taking a multitude of fifteen units some number of times. That is to say: the multitude of fifteen units does not “measure” forty units. Both fifteen units and forty units are, however, measured by five units; and five units, which is the eighth part of forty units (the part obtained by dividing forty units into eight equal parts), is also the third part of fifteen units, so fifteen units is three of the eighth parts of forty units. This is to say that fifteen units, which is not a “part” of forty units, *is* “parts” of it. Indeed, a smaller number, if it is not part of a larger number, must be parts of it, since both numbers must—just because they are numbers—be measured by the unit.

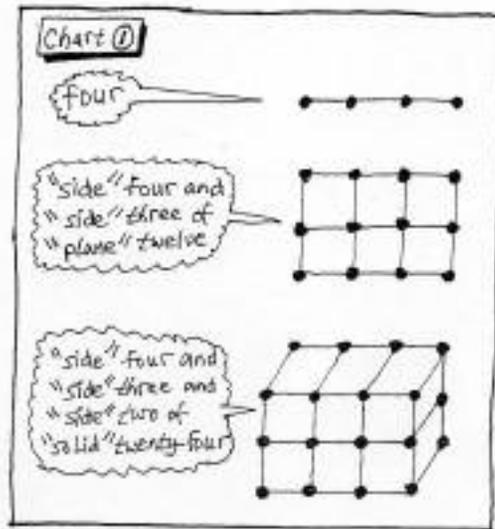
After that, Euclid goes on to tell us about sorts of numbers, and the ways in which they are related, numbers such as those that are “even,” or “odd,” or “even-times-even,” or “even-times-odd,” or “odd-times-odd,” or “prime,” and so on.

Numbers are thus of interest not only with respect to their relations of size, but also with respect to their relations of kind. Even and odd are distinguished by divisibility into two equal parts (or by odd’s differing from even by a unit), and the more complex terms here are distinguished with respect to “measuring by a number according to another number”—which nowadays would be stated thus: “dividing by an integer so as to yield another integer as quotient without a remainder.”

Euclid thus defines some kinds of numbers by employing notions such as being greater than or equal to or having a certain difference in size. However, he also sorts numbers into kinds by referring to shape. What might lead to doing that?

We might use a dot to represent a unit, and then represent a number by dots in a line. For example, four would be represented as it is at the top of Chart 1. We might then rep-

resent in a shape the number produced by taking a number some number of times. For example, the number produced when four is taken three times would be represented as it is in the middle of the chart, where we see that the numbers four and three are “sides” of the “plane” number twelve; and similarly, six and four are “sides” of the “plane” number twenty-four. But the number twenty-four can also be represented as a solid, as it is at the bottom of the chart: the numbers four and three and two are “sides” of the “solid” number twenty-four.



So Euclid speaks of numbers that are “sides,” “plane,” and “solid,” and also of the multiplication of a number by a number, and then of the sorts of numerical products of such multiplication—such as the numbers called “square” and “cube,” whose factors are called “sides.”

But he does *not* represent numbers by dots in lines. To use dots would force us to pick this or that number. To signify not this or that number, but rather any number at all, it is more convenient to use bare lines, without indicating how many dots are carried on each line, although it might be confusing that we then also have to represent by a line the unit.

On the other hand, since the unit does have a ratio to any number, as does any number to any other number, and since the product of multiplying any number by any other number is also a number, it might make everything more convenient if we represent as lines not just the numbers that are “sides,” as well as any “side” that is a unit, but also all the figurate numbers that we can produce. After all, though the sort of number called a “square” number may be somehow different from the sort of number called a “cube,” it is nonetheless true that any “square” number has a ratio to any “cube” number (since *any* number has a ratio to any *other* number), whereas a square cannot have a ratio to a cube (since they are not magnitudes of the same kind). If we go along with speaking in this way, however, we must take care to keep in mind that when it is numbers that are called “sides” or “squares” or “cubes” we are not engaged in speech about figures; we are using figures of speech. Several of the books of Euclid’s *Elements* deal with what would nowadays be called “number theory” rather than “geometry.”

The numbers in the classical sense—the multitudes two, three, four, and so on—tell us the result of a count, and nowadays we often speak as if those numbers, the counting-numbers, are merely some of the items contained in an expanded system that we call “the *real* numbers.”

Whenever we put numbers of things together we get some number of things, and whenever we take a number of things a number of times we get some number of things; but only sometimes can we take a given number of things away from a given number of things, and even when we can, only sometimes can we get some number of things by doing so. For example, we cannot take seven things away from five things; and although we can take six things away from seven things, we will not have a *number* of things left if we do, but only a single thing. It is also only sometimes that we can divide a given number of things into a given number of equal parts. A multitude of ten things can be divided into five equal parts, but a multitude of eight things cannot. In dividing a many-

ness, there is constraint from which we are free in dividing a muchness.

We often measure a muchness to say just how much of it there is compared with some other muchness: having chosen some unit of muchness, we count how many times such a unit must be taken in order to be equal to the muchness that we are measuring. A line, for example, is a kind of a muchness: there is more line in a longer line than in a shorter line. If two lines are not of equal size, one of them is greater than the other. Lines are magnitudes. If we measure them, we obtain multitudes that tell their sizes—their lengths. But magnitudes are not multitudes. Whereas the size of a magnitude has to do with how much of it there is, the size of a multitude has to do with how many units it is composed of.

Not only magnitudes, but *multitudes* as well are often measured: we measure these by using other multitudes to measure them. Thus, we say that twelve is six taken two times, or is twice six. We say, moreover, that eight is twice the third part of twelve, or—to abbreviate—that eight is two-thirds of twelve. “Two-thirds” does not name a multitude (unless we say that we are merely treating a third part of twelve as a new unit) but “two-thirds” is nonetheless a numerical expression. If we therefore begin to treat it like the name of a number, we are on the road to devising a system of fractions. “Two-thirds” of something is *less* than one such something, but it is not *fewer*. Two-thirds *are* fewer than three-thirds—but to say this, we must have broken the unit into three new and smaller units. The numerator of a fraction is a number or else it is “one.” The denominator is also a number: it is the number of parts that result from a division (although we will allow a denominator as well as a numerator to be “one.”) With a fraction, we divide into equal parts and then we count them. A fraction is not itself a number (a multitude), but it may be the counterpart of a number. “Twelve-sixths” is not the number “two”; but it *is* a counterpart, among the fractions, of the item “two” among the numbers. Another counterpart of the number “two” is the fraction

“eight-fourths.” Those counterparts are really the same counterpart differently expressed, since they are both equal to the fraction whose numerator is the number “two” and whose denominator is “one.”

As with division, so with subtraction. In taking things away, we may get a numerical expression that we can treat like a number in certain respects. When we have twelve horses and eight of them are taken away, there is a remainder of four horses. But if eleven are taken away, then there are not horses left; only a single horse remains. And if twelve horses are taken away, then not even a single horse is left: none remain. But twelve horses can be compared with a single horse: the twelve are to the one as twenty-four are to two. “Twenty-four” and “two” are numbers, and “twelve” is also a number—so “one” is like a number in being comparable with a number as a number is. Moreover, “one” can be added to a number as a number can be, and “one” is sometimes what we reply when asked how many of some kind of thing there are. “One” is therefore a numerical word even if it is not a number. But in that case, so is “none”: if twelve horses are taken away, leaving not a single horse, and we are asked how many remain, we can then say “none.” Now suppose that Farmer Brown owes fourteen horses to Farmer Gray, but has in his possession only ten horses. Farmer Gray takes away the ten. How many horses does Farmer Brown now own? None. But he might be said in some sense to own even less than none, for he still owes four. If we did say that, then we would have to say that he owns four *less* than no horses, or that fourteen less than ten is four less than none. But then we would be counting, not horses-*owned*, but horses-*either-owned-or-owed*. We would be letting the payment of four-horses-owed count as wealth equal to no-horses-either-owed-or-owned. Again, as with the fractions, we have numerical expressions that we are beginning to treat like numbers. We are on the road to devising a system of so-called “rational numbers,” which includes “negative” items as well as such non-negative items as “zero” and “one” in addition to fractions (which

include the *counterparts* of those multitudes that were first called “numbers”—namely, two, three, four, and so on).

We must say “counterparts” because the multitude “three,” for example, while it corresponds to, is not the same as “the positive fraction nine-thirds” or “the positive integer three.” We have not brought about an *expansion* of a number system by merely introducing *additional* new items alongside old ones; rather, we have made a new system which contains some items that *correspond* to the items in the old system, but that *differ* from the old items in being items of the very same *new* sort as are the new items that do *not* have correspondents in the old system. For example: although the number “three” is a multitude, “positive three” is not a multitude; it is an item defined by its place in a system where, among other things, “positive six” takes the place belonging to the item that is the outcome of such operations as multiplying “negative two” by “negative three.”

We have just taken a look at part of the road that leads from numbers in the classical sense to what we nowadays are used to thinking of as numbers. We have looked at some steps on the road to the so-called “rational numbers,” the so-called numbers that have something to do with ratios. But the system of what we nowadays are used to thinking of as numbers is the system of “real numbers,” only some of which are items that have counterparts among the items in the system of “rational numbers.” The system of what we nowadays are used to thinking of as numbers is replete with “irrational numbers,” some of them “algebraic” (such as “the square root of two”) and some of them “transcendental” (such as “pi”).

Classically, numbers are multitudes, and there may be some number-ratio that is the same as the relation in size of some one magnitude to another. But there does not have to be a number-ratio that is the same. If there did have to be such a ratio of numbers for every ratio of magnitudes, then there would not have been a reason for devising a system of “real numbers.” The reason for taking the step from “rational

numbers” to “real numbers” has to do with the difference between *multitudes* and *magnitudes*.

The difference between how we can speak about multitudes (that is, numbers in the classical sense) and how we can speak about magnitudes classically manifests itself in the statements with which Euclid begins his treatment of magnitudes in the Fifth Book of the *Elements* and his treatment afterwards of numbers in the Seventh Book.

Although Euclid says what number is, he does not say what magnitude is. Examples of magnitudes can be found in his propositions, however. After the Fifth Book of the *Elements* demonstrates many propositions about the ratios of magnitudes as such, these are used by later books of the *Elements* to demonstrate propositions about such magnitudes as straight lines, triangles, rectangles, circles, pyramids, cubes, and spheres. Such magnitudes correspond to what we nowadays call lengths, areas, and volumes. Weight is yet another sort of magnitude. Though weights are not mentioned in the *Elements*, what Euclid says there about magnitudes generally is applied by Archimedes to weights in particular.

At the beginning of the Fifth Book, Euclid defines ratio for magnitudes; but at the beginning of the Seventh Book he does not define ratio for numbers. There in the Fifth Book he also defines magnitudes’ being in the same ratio, and only after doing that does he define magnitudes’ being proportional. But here in the Seventh Book he does not define numbers’ being in the same ratio: he goes directly into defining numbers’ being proportional, which he needs to do in order to define the similarity of numbers that are of the sorts called “plane” or “solid.”

The Greek term translated as “proportional” is *analogon*. After a prefix (*ana*) meaning “up; again,” the term contains a form of the word *logos*. This is the Greek term translated as “ratio.” *Logos* is derived from the same root from which we get “collect” (which is what the root means); and in most contexts, it can be translated as “speech” or as “reason.” Proportionality is, in Greek, *analogia* (from which we get

“analogy”), a condition in which terms that are different may be said to carry or hold up again the same articulable relationship.

Later in the *Elements*, in the enunciation of the fifth proposition of the Tenth Book, for example, Euclid does use the term “same ratio” in speaking of ratios that are numerical. In the definitions with which the Tenth Book begins, Euclid says that lines which have no numerical ratios to a given straight line are called “*logos-less*”—that is, *alogoi*. This Greek term for lines lacking any articulable ratios to a given line is translated (through the Latin) as “irrational.”

Such a line and the given line to which it is referred cannot both be measured by the same unit, no matter how small a unit we may use to try to measure them together. They are “without a measurement together”—that is, *asymmetra*. This Greek term is translated (through the Latin) as “incommensurable.” Because magnitudes of the same kind can be incommensurable, magnitudes are radically different from multiples, and so we must speak of them differently.

Part Three

Muchness Not Related Like Manyness: Incommensurability

Let us turn now to the classic example of incommensurability: let us consider the relation between the side of a square and its diagonal.

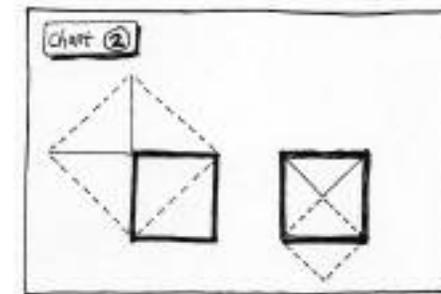
If a square’s diagonal were in fact commensurable with its side, then the ratio of the diagonal to the side would be the ratio of some number to some other number. But it cannot be that. Why not? Because if it were, then it would have two incompatible properties: one property belongs to *any* ratio whatever which *a number has to a number*, and the other property follows from that *particular* ratio which *a square’s diagonal has to its side*. Let us convince ourselves that there is such a contradiction.

First, let us look at that property which belongs to any ratio whatever that is numerical. It is this: any numerical ratio whatever must either be in lowest terms already, or be

reducible to them eventually. Consider, for example, the ratio that thirty has to seventy. Those two numbers have in common the factors two and five; so, dividing each of them by ten, we see that the ratio that thirty has to seventy is the same as the ratio that three has to seven. The ratio of three to seven is the ratio of thirty to seventy reduced to lowest terms. If a ratio of one number to another number were not reducible to lowest terms, then the two original numbers would have to contain an endless supply of common factors, which is impossible: any number must sooner or later run out of factors if you keep canceling them out by division.

This property of every ratio that a number has to a number cannot belong to the special ratio that a square’s diagonal has to its side. To see why this is so, consider a given square. If you make a new square using as the side of the new square the diagonal of the given square, then the new square will be double the size of the original square.

Chart 2 shows that doubling. In the left-hand portion of the chart, the diagonal of a square divides it into two triangles, and it takes *four* of these triangles to fill up a new square which has as its own side that diagonal. In the right-hand portion, the chart also shows what happens if we now make another new square, using as this newest square’s own side a *half of the diagonal* of the original square. The original square, now being itself a square on the newest square’s diagonal, will itself be double the size of this newest square—since this newest square is made up of two triangles, and it takes four of these triangles to make up the original square.



Let us suppose that it were in fact possible to divide a square's diagonal as well as its side into pieces that are all of the same size. Let us count the pieces and say that K is the number of pieces into which we have divided the diagonal, and that L is the number of pieces into which we have divided the side. By taking as a unit any of those equal pieces into which we have supposedly divided the diagonal and the side, we have measured the two lines together: the ratio which the number K has to the number L would be the same as the ratio which the square's diagonal has to its side.

Now, disregarding for a moment just what ratio the number K has to the number L , but considering only that it is supposed to be a ratio of numbers—which, as such, must be reducible to lowest terms—we can say that there would have to be two numbers (let us call them P and Q) such that (1) P and Q do not have a single factor in common and (2) the ratio that P has to Q is the same ratio that K has to L . Since the ratio that K has to L is supposed to be a ratio of numbers, there cannot be any such pair of numbers as K and L unless there is also such a pair of numbers as P and Q . (If the ratio of K to L is already in lowest terms, then we will just let K and L themselves be called by the names P and Q .) So now we have P being the number of equal pieces into which the square's diagonal is divided, and Q being the number of such pieces into which the square's side is divided. And now we will see that there just cannot be any such numbers as P and Q because this pair of numbers would have to satisfy contradictory requirements.

The first requirement results from the necessities of *numbers*; the second, from the consequences of *configuring lines*. Because (as has been said) the ratio of P to Q is a ratio of numbers reduced to lowest terms, P and Q are numbers that cannot have a single common factor; but (as will be shown) because the ratio of P to Q is the same ratio that a square's diagonal has to its side, P and Q must both be numbers divisible by the number two. That is to say, P and Q are such that they cannot both be divisible by the same number, and yet

they also must both be divisible by the number two—a direct contradiction.

Now we must see just why the latter claim is true. We ask: why is it, that if any pair of numbers are in that special ratio which a square's diagonal has to its side, then both numbers must be divisible by the number two?

In order to see why, we must first take note of another fact about numbers: if some number has been multiplied by itself and the product is an even number, then the number that was multiplied by itself must *itself* have been a number that is even. (The reason is simple. A number cannot be both odd and even—it must be one or the other—and when two even numbers are multiplied together, then the number that is the product is also even; however, when two odd numbers are multiplied together, then the number that is the product is not even but odd.) That said, we are now ready to see why it is that if any two numbers are in that special ratio which a square's diagonal has to its side, then they must have the factor two in common.

The numbers P and Q have the factor two in common because both of them must be even. They must both be even because each of them when multiplied by itself will give a product that must be double another number—and therefore each must itself be double some number. Why?

Look at the left-hand portion of Chart 3, and consider why P must be an even number. Because the square on the diagonal is double the square on the side, the number produced when P is taken P times is double the number produced when Q is taken Q times. Hence (since any number which is double some other number must itself be even) the multiplication of P by itself produces an even number, and hence the number P itself must be even.

And now look at the right-hand portion of the chart, and consider why Q also must be an even number. Since it has already been shown that P must be an even number, it follows that there must be another number that is half of P —call this number H . But because the square on the side is double the

whatever unit measures either one of them will not measure the other one too. By a unit's "measuring" a magnitude, we mean that when the unit is multiplied (that is, when it is taken some number of times), it can equal that magnitude. Therefore, when magnitudes of the same kind are incommensurable, one of them has to the other a ratio that is not the same as any numerical ratio.

The demonstrable existence of incommensurability means that comparisons of size can only be approximated by using numbers. We can, to be sure, be ever more exact—even as exact as we please—but if we wish to speak with absolute exactness, numbers fail us.

For the ancient Pythagoreans, who were the first to conceive of the world as thoroughly mathematical, knowledge of the world was knowledge of numerical relationships. In the world stretching out around us, the Pythagoreans saw correlations between shapes and numbers, such as we encountered when we considered kinds of numbers, "square" numbers and "cube" numbers, for example.

The Pythagoreans noted also that the movements of the heavenly bodies take place in cycles. Their changes in position, and their returns to the same configuration, have rhythms related by recurring numbers that we get by watching the skies and counting the times. With numbers, the world goes *round*. We are surrounded by a cosmos. (*Cosmos* is the Greek word for a "beautiful adornment.") The beauty on high appears to us down here in numbers.

But even the qualitative features of the world of nature show a wondrous correlation with numbers: numbers make the world sing. It is not merely that rhythm is numerical, it is that tone or at least pitch, is too. If a string stretched by a weight is plucked, it gives off some sound. The pitch of the sound will be lowered as the string is lengthened. If another string is plucked (a string of the same material and thickness, and stretched by the same weight) then the longer the string the lower will be the pitch of the sound produced by plucking it. Now, what set the Pythagoreans thinking was the rela-

tion between numbers and harmony. (*Harmonia* is the Greek word for "the condition in which one thing fits another"; a word from carpentry thus is used to describe music.) When the lengths of two strings of the sort just mentioned are adjusted so that one of them has to the other the same ratio that one small number has to another, or to the unit, then there is music. With the Pythagoreans' mathematicization of music, mathematical physics begins.

Thus not only were the sights on high seen to be an expression of mathematical relationships, so were the sounds down here that enter our souls and powerfully move what lies deep down within us. Thinking that nature is a display of numerical relationships, and that human souls are gotten into order by attending to those relationships, the Pythagoreans formed societies that sought to shape the thinking of the political societies of their time by being the givers of their laws. It is said that when the discovery of incommensurability was first revealed to outsiders, thus making public the insufficiency of number, the man who thus had undermined the Pythagorean enterprise was murdered.

But never mind the whole wide world; even the relationships of size in mere geometrical figures cannot be understood simply in terms of multitudes. If geometry could in fact be simply arithmeticized; if we could just measure lines together, getting numbers which we could then just multiply together, and thus simply express as equations all the relationships that we have to handle, then much in Euclid, in Apollonius, and in the other classical mathematicians that is difficult to handle could be handled much more easily—and Descartes in the seventeenth century would not have had to undertake a radical transformation of geometry. Instead of manipulating equations, however, we must in the study of classical mathematics learn to deal with non-numerical ratios of magnitudes, and with boxes that are built from lines devised to exhibit those ratios.

However, before we can freely deal with ratios, we must learn what can be meant by calling two ratios the same—even

when they are not the same as some numerical ratio. Euclid tells us that in the Fifth Book of the *Elements*. That is long before his discussion of number, which does not take place until the Seventh Book, thus raising a question about what kind of a teacher he is: after all, is it not a principle of good teaching that questions should be raised before answers are presented?

Incommensurability is responsible for the difficulty of Euclid's definition of sameness of ratio for magnitudes, as well as for the difficulty that modern readers encounter in the classical presentation of relationships of size generally. Let us now look at that definition.

Part Four

Muchness Related After All: Euclid's Definition of Same Ratio

We insist that even when two magnitudes are incommensurable, their ratio can be the same as the ratio of two other incommensurable magnitudes. For example, we insist that the ratio which one square's diagonal has to its side is the same ratio which another square's diagonal has to its own side. Sameness of ratio for magnitudes is therefore not to be defined in terms of measuring magnitudes: we could so define it only if we could divide each of the magnitudes into pieces, each equal to some magnitude that is small enough to be a suitable unit, and then, when we counted up all of those small equal pieces, we found no left-over unaccounted-for even-smaller piece of either of the magnitudes; but we cannot do that with magnitudes that are incommensurable. Although we can say that two ratios of magnitudes are the same as each other if they are both the same as the same ratio of numbers, we will not say that they are the same as each other only if they are both the same as the same ratio of numbers.

What, then, is to be said instead? According to Euclid, this: "Magnitudes are said to be in the same ratio, the first to the second and the third to the fourth, when, if any equimultiples whatever be taken of the first and third, and any equi-

multiples whatever of the second and fourth, the former equimultiples alike exceed, are alike equal to, or alike fall short of, the latter equimultiples respectively taken in corresponding order."

When Euclid says for ratios of magnitudes, what sameness is, it is very difficult at first to understand just what he means. While his definition may be the proper departure point on a road that we need to travel, for a beginner it seems to constitute a locked gate.

A key to open that locked gate, however, is this fact: while there is *no* ratio of numbers that is *the same as* a ratio of incommensurable straight lines, nonetheless *every* ratio of numbers is *either greater or less than* such a ratio of straight lines. That is to say: whatever ratio of numbers we may take, we can decide whether the ratio of one line to another is greater or less than it.

For example, we might ask: which is greater—the ratio of the one line to the other, or the ratio of seven to twelve? We would divide the second line into twelve equal parts, and then take one of those pieces in order to measure the first line. Suppose that this first line, although it may be longer than six of these pieces put together, turns out to be shorter than seven of them. We would conclude that the first line has to the other line a ratio that is less than the numerical ratio of seven to twelve.

Let us consider the ratio of lines that has been of such concern to us, the ratio which a square's diagonal has to its side. And let us take the following ratio of numbers: the ratio which that three has to two. As we have seen, the ratio of lines that we are considering cannot be the same as *any* ratio of numbers whatever; so it must be either greater or else less than that ratio of numbers which we have taken. Which is it?

It *must* be less—because if it were greater (that is: if the diagonal were greater than three-halves of the side), then the diagonal taken two times would have to be greater than the side taken three times. But if that were so, then a new square whose side was the diagonal of the original square would

have to be more than double the original square, as is shown in Chart 4.



Having thus shown that the ratio of a square's diagonal to its side is less than the ratio which three has to two, we could in like manner also show that the ratio of those two lines is greater than the ratio that, say, four has to three. And we could go on showing, for the ratio of *any* pair of numbers which we choose, that the ratio of a square's diagonal to its side is greater or is less than the ratio of the pair of numbers chosen.

Indeed, although no ratio of numbers is the same as the ratio of a square's diagonal to its side, we can nonetheless confine this ratio of lines as closely as we please by using pairs

of ratios of numbers, as follows. (The results are shown in Chart 5.) Let us divide the square's side into ten equal parts, and take such a tenth part as the unit; then the diagonal will be longer than fourteen of these units but shorter than fifteen of them; then let us consider what we get when we take as the unit the side's hundredth part, and then its thousandth. We can go on and on in that way, eventually reaching numbers large enough to give us a pair of number-ratios that differ from each other as little as we please; we can show it to be *greater* than the ever-so-slightly-smaller number-ratio, and smaller than the ever-so-slightly-greater number-ratio. So, although the ratio of a square's diagonal to its side cannot be the same as any number-ratio, it can be confined between a pair of number-ratios that differ from each other as little as we please.

Chart 5

divide square's side into	10 equal pieces	100 equal pieces	1,000 equal pieces
so the unit will be	a 10th part of the side	a 100th part of the side	a 1,000th part of the side
then the diagonal will be shorter than	15 units	142 units	1,415 units
but longer than	14 units	141 units	1,414 units

less than ratio of 15 to 10

less than ratio of 142 to 100

less than ratio of 1,415 to 1,000

greater than ratio of 1,414 to 1,000

greater than ratio of 141 to 100

greater than ratio of 14 to 10

The question that we want to answer, however, is not when it is that ratios of magnitudes are almost the same, but rather when it is that they are the *very* same—even when they are not both the same as the same ratio of numbers. The key to the answer, restating what we said, is this: given any ratio of magnitudes, any ratio of *numbers* that is *not the same* as it must be *either greater* than it *or smaller*.

That gives us the following answer: given *two* ratios of *magnitudes*, they will be the same as *each other* (whether or not they are the same as some ratio of *numbers*) whenever it is true that—taking *any* ratio of *numbers* whatever—if *this ratio of numbers* is *greater* than the *one* ratio of *magnitudes*, then it is *also greater* than the *other*; and if it is *less* than the one, then it is *also less* than the *other*.

Euclid's definition seems much more complicated than that, however, because it emphasizes "equimultiples." Why does it do that? Because that makes it simpler to compare ratios of magnitudes with ratios of numbers: taking multiples is a way to make the comparisons *without* performing any divisions. We could in fact get a definition by using division, but the definition would be clumsier if we did.

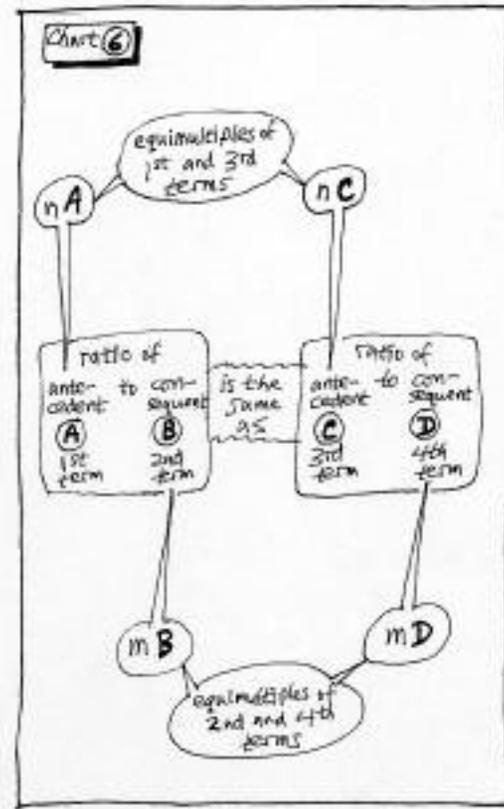
Let us put the notion of *equi*-multiples aside for a moment, and consider multiples *simply*, in contradistinction to divisions into parts.

Suppose, for example, that the ratio of some magnitude (A) to another magnitude (B) is greater than the ratio of nine to seven but less than the ratio of ten to seven. This means that if we divide B into seven equal parts and we use as a unit (for trying to measure A) one of those seventh parts of B, then A will turn out to be greater than nine of those parts of B, but less than ten of them. If A and B happen to be incommensurable, then, no matter how many equal parts into which we divide B—that is, no matter how small we make the B-measuring unit with which we try to measure A, we will find that we cannot divide A into pieces of that size without having a smaller piece left over.

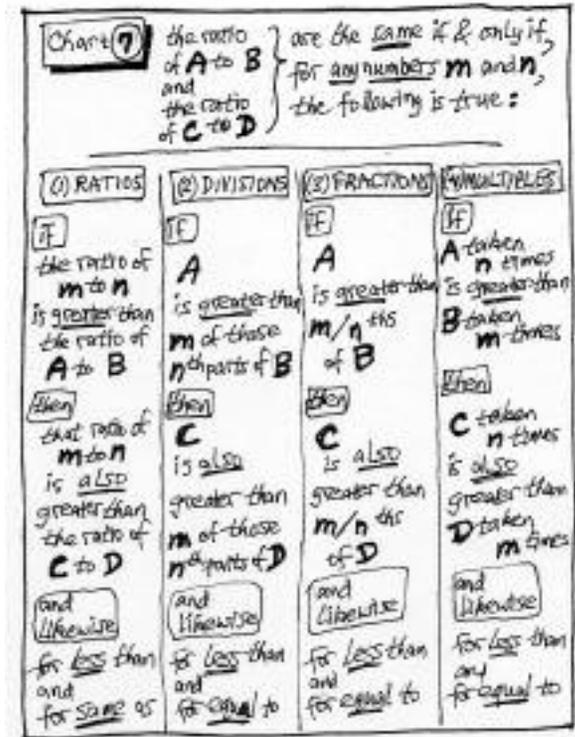
So, in comparing the ratio of A to B with the numerical ratio of some multitude *m* to some multitude *n*, it is less awkward to speak of A-taken-*n*-times and of B-taken-*m*-times than to speak of the little piece of A that may be left over when we divide A into *m* pieces that are each equal to the *n*th part of B—or, in other words, the little piece that may be left over in A, no matter how small are the equal parts into which we divide B. To make the requisite comparisons, we *must* consider *multiples* of magnitudes, but we *need not* consider *parts* of them; we *have to multiply*, but we *do not* also have to *divide*. Instead of trying to measure magnitudes together, by dividing them and counting the parts, we can speak merely of *multiples of the magnitudes*.

What we have just now seen is this: to say that the ratio of magnitude A to magnitude B is either the same as the numerical ratio of *m* to *n*, or is greater or less than it, is to say that A-taken-*n*-times is either equal to B-taken-*m*-times, or is greater or less than it. Let us take that and put it together with what we saw earlier—which was this: to say that the ratio of A to B is the same as the ratio of C to D is to say that whatever numerical ratio you may take (say, of *m* to *n*) this ratio of numbers will not be the same as, or greater than, or less than one of the ratios of magnitudes unless it is likewise so with respect to the other one.

All that remains is for multiples of a certain sort to be brought in—namely, "*equi*-multiples." Equimultiples of two magnitudes are two other magnitudes that are obtained by multiplying the two original magnitudes an equal number of times. Here, as Chart 5 shows, both ratios' antecedent terms (namely, A and C) are each taken *n* times, and their consequent terms (namely, B and D) are each taken *m* times. In other words, equimultiples (*nA* and *nC*) are taken of the *first* and *third* magnitudes (A and C); and *also* equimultiples (*mB* and *mD*) are taken of the *second* and *fourth* magnitudes (B and D). And *m* and *n* are any numbers at all: we are interested in all the equimultiples of the first and third magnitudes, and also of the second and fourth ones.



Now at last we are in a position to see that Euclid's definition is not so bewildering as might have seemed at first glance. There are several ways of saying what we can do with what we have seen. In abbreviated form, they are exhibited in Chart 7.



Now in interpreting the chart, just remember that whenever we say "IF something, THEN something else," that is equivalent to saying "NOT something UNLESS something else." So, the ways laid out in the chart are these:

First way: Staying with ratios, we can say that two ratios of magnitudes (call them the ratios of A to B , and of C to D) are the same if, and only if, whatever numerical ratio we may take (say the ratio that some number called m has to some other number called n) the following is true: that numerical ratio which m has to n will not be the same as, or greater than, or less than one of the two ratios of magnitudes, unless it is likewise so with respect to the other one.

Second way: We might want to get more explicit by making divisions into equal parts—that is, divide B into n equal parts and do the same to D (these, B and D , being the conse-

quent terms of the two ratios). Then the ratios would be the same if the following is true: antecedent term A will not be greater than, or less than, or equal to m of those n^{th} parts of its consequent term B unless the other antecedent term C is likewise so with respect to that same number m of those n^{th} parts of its own consequent term D.

Third way: Someone who was willing to do all that (namely, willing to divide magnitudes into equal parts and then count them up and compare sizes) might insist on using fractions to restate that as follows: A will not be greater than, or less than, or equal to m/n^{ths} of B unless C is likewise so with respect to that same fraction (m/n^{ths}) of D.

Fourth, and final way: Rather than getting into the complicated business of dividing, counting divisions, and comparing sizes, even though we can briefly restate it all by using fractions, we might simply take multiples alone; and then we would say that the ratios are the same if the following is true: A-taken- n -times will not be greater than, or less than, or equal to B-taken- m -times unless C-taken-exactly-as-many-times-as-A is likewise so with respect to D-taken-exactly-as-many-times-as-B.

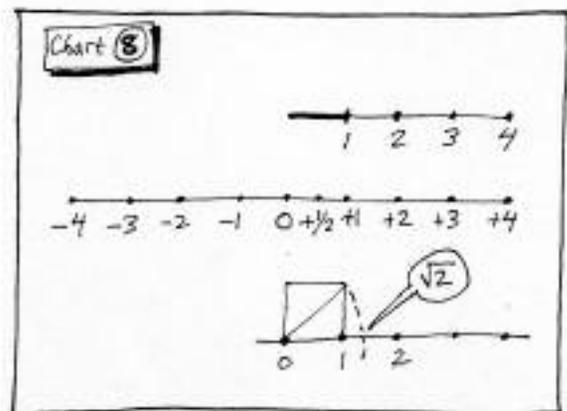
Those are several ways to determine, despite incommensurability, when we may say that the ratio of one magnitude to another is the same as the ratio of a third magnitude to a fourth one. The first way formulates our initial insight, and the final way formulates Euclid's definition of same ratio for magnitudes. Euclid's definition, like the others, covers all ratios of magnitudes, whether or not they are the same as ratios of multitudes; but, unlike the others, it manages to do so by speaking simply of multitudes of magnitudes.

Magnitudes and multitudes *are not the same* as each other, nor is either of them the same as ratios of them, but magnitudes and multitudes and their ratios are all, in a way, alike. How?

In the first place, any two magnitudes of the same kind are equal, or else one of them is greater than the other; and likewise, any two multitudes are equal, or else one of them is

greater than the other. Moreover, any magnitude has a ratio to any other magnitude of the same kind, just as any multitude has a ratio to another multitude. Finally, any two ratios (whatever it may be that they are ratios of, whether magnitudes or multitudes) are the same as each other, or else one of them is greater than the other. This last fact supplies the insight that enabled Euclid to define same ratio for magnitudes regardless of whether the magnitudes are commensurable.

Euclid's definition can help us to understand what enabled Dedekind several thousand years afterward to define "the real numbers" in terms of "the rational numbers." Dedekind found himself in the following situation. In modern times, after proportions containing magnitudes and multitudes had given way to equations containing "real numbers," there was still some difficulty in saying just what "real numbers" were. To say much about the matter, it seemed necessary to refer not only to multitudes to which we come by counting (that is, numbers in the strict sense) but also to magnitudes that we visualize (namely, lines). If we place the counting numbers along a line (as in the top portion of Chart 8), it is then clear not only where to put all the "rational numbers," including those which are fractional and those which are non-positive (as in the middle portion of the chart)—but also where to put such "irrationals" as "the square root of two." To put "the square root of two" in its place, for example, (as in the bottom portion of the chart) just erect a square that has a side that is a unit long, and swing its diagonal down.



But if “the square root of two” is a number, it should be definable without appealing to visualization. After all, we need not visualize in order to count, or even to go on to think up fractions and negatives. That thought troubled Dedekind as he taught calculus in the nineteenth century, relying upon appeals to a number-line. He finally saw how to treat “real numbers” (the modern counterpart of Euclidean ratios of linear magnitudes) in terms of collections of “rational numbers” (the modern counterpart of Euclidean ratios of numbers), thus making it more plausible to speak of the “real numbers” as really numbers.

Reading Dedekind is one way to think about what it means to rely on a system of “real numbers.” Unless we do think through the meaning of relying on a system of real numbers, many of the things we take for granted in the modern world are without foundation.

But does Dedekind’s work simply represent progress beyond Euclid’s, or are very important differences covered over by the important similarities between what Dedekind had his mind on and Euclid his?

Ratios relate multitudes, or magnitudes. *Ratios* (like *multitudes*) have *homogeneity*, and also (like *magnitudes*) have *continuity*; but magnitudes do *not* have homogeneity, and multitudes do *not* have continuity. Ratios themselves are

therefore *not* things of the very same sort as the things that they relate.

Ratios are not things of the same sort as quotients either. To be sure, ratios are like quotients in that they are quantitative. Like any two magnitudes of the same sort, or any two multitudes—or any two quotients—any two ratios (whether of magnitudes or of multitudes) have an order of size. But quotients, unlike ratios, are like what they are quotients *of*: a quotient is itself a “real number” that is obtained through the operation of dividing a “real number” by a “real number”; if A and B are “real numbers,” then the quotient A/B is a “real number” also.

Yet, although ratios and “real numbers” are things of different sorts, they do have a similarity—and not merely that they are both quantitative. Although the magnitudes are not homogeneous, and the multitudes are not continuous, the ratios are both homogeneous and continuous—and so also are “the real numbers.”

Part Five

Mathematics and the Modern Mind

In all this, we have been considering the classical handling of numbers and lines in its difference from modern notions which conflate the two, but we have not considered what it was that led to the conflation. That is, indeed, an immensely important matter, but it is too long a story to be told now. To tell that tangled tale, one must traverse much of the road that constitutes the Mathematics Tutorial at St. John’s College. Along that arduous road, it is easy to be overwhelmed, however, and thus to lose sight of why it is worth our while to traverse it at all—so perhaps at least something should be said about it now.

Euclid’s *Elements* prepares us for a higher study in geometry, the *Conics* of Apollonius. This was, for almost two-and-a-half millenia, the classic text on the curves which—following the innovative terminology of Apollonius—came to be called the “parabola,” the “hyperbola,” and the “ellipse.”

Apollonius' conic sections were lines first obtained upon a plane by cutting a cone in various ways; they were then characterized by relative sizes and shapes of certain boxes formed by associated straight lines standing in certain ratios. After Descartes, although the names of those curves persisted, and they continued to be called collectively "the conic sections," they nonetheless eventually ceased to be studied as such: they came to be studied algebraically. But although Descartes' *Geometry* may be in fact what it has been called—the greatest single step in the progress of the exact sciences—no one could clearly see it as that without studying Apollonius, for Cartesian mathematics has shaped the world we live in and shapes our minds as well.

Apollonius tells his tale obliquely. He does not give us questions, but rather gives us only answers that are too hard to sort out and remember unless we ourselves figure out what questions to ask. About Apollonius as a teacher we must ask whether his work is informed by wisdom and benevolence. Descartes did not think so. In his own *Geometry*, and in his sketch of rules for giving direction to the native wit, Descartes found fault with the ancient mathematicians.

Descartes severely criticizes them—for being show-offs. He says that they made analyses in the course of figuring things out, but then, instead of being helpful teachers who show their students how to do what they themselves had done, they behaved like builders who get rid of the scaffolding that has made construction possible. Thus they sought to be admired for conjuring up one spectacular thing to look at after another, without a sign of how they might have found and put together what they present.

Descartes also suggests, however, that they did not fully know what they were doing. They did not see that what they had could be a universal method. They operated differently for different sorts of materials because they treated materials for operation as simply objects to be viewed. Hence they learned haphazardly, rather than methodically, and therefore they did not learn much. Mathematics for them was a matter

of wonderful spectacle rather than material for methodical operation. The characteristic activity of the ancient mathematicians was the presentation of theorems, not the transmission and application of the ability to solve problems.

They had not discovered the first and most important thing to be discovered: the significance of discovery. They had not discovered the power that leads to discovery and the power that comes from discovery. They were not aware that the first tools to build are tools for making tools. They were too clever to be properly simple, and too simple to be truly clever. They were blinded by a petty ambition. Too overcome by their ambition, they could not be ambitious on the greatest scale.

Were the ancient mathematicians as teachers guilty of the charges set out in that Cartesian critique? Were they guilty of the obtuseness of which Descartes in his *Geometry* accused them—were they guilty of the desultory fooling-around and disingenuous showing-off of which Descartes had accused them earlier, in his *Rules*? You cannot know unless you study them.

In any case, the study of Euclid and Apollonius gives access to the sources of the tremendous transformation in thought whose outcome has been the mathematicization of the world around us and the primacy of mathematical physics in the life of the mind. Scientific technology and technological science have depended upon a transformation in mathematics which made it possible for the sciences as such to be mathematicized, so that the exact sciences became knowledge par excellence. The modern project for mastering nature has relied upon the use of equations, often represented by graphs, to solve problems. When the equation replaced the proportion as the heart of mathematics, and geometric theorem-demonstration lost its primacy to algebraic problem-solving, an immense power was generated. It was because of this that Descartes' *Geometry* received that accolade of being called the greatest single step in the progress of the exact sciences. To determine whether it was indeed such a step, we need to

know what it was a step from as well as what it was a step toward. We cannot understand what Descartes did to transform mathematics unless we understand what it was that underwent the transformation. By studying classical mathematics on its own terms, we prepare ourselves to consider Descartes' critique of classical mathematics and his transformation not only of mathematics but of the world of learning generally—and therewith his work in transforming the whole wide world.

It is in the study of Apollonius on the conic sections that the modern reader who has been properly prepared by reading Euclid can most easily see both the achievement of classical mathematics and the difficulty that led Descartes and his followers to turn away from it.

It all has to do with ratio, and with notions of number and of magnitude. For Apollonius, as for Euclid before him, the handling of ratios is founded upon a certain view of the relation between numbers and magnitudes. When Descartes made his new beginning, almost two millenia later, he said that the ancients were handicapped by their having a scruple against using the terms of arithmetic in geometry. Descartes attributed this to their not seeing clearly enough the relation between the two mathematical sciences. Before modern readers can appreciate why Descartes wanted to overcome the scruple, and what he saw that enabled him to do it, they must be clear about just what that scruple was. Readers must, at least for a while, make themselves at home in a world where *how-much* and *how-many* are kept distinct, a world which gives an account of shapes in terms of geometric proportions rather than in terms of the equations of algebra. For a while, readers must stop saying "AB-squared," and must speak instead of "the square arising from the line AB"; they must learn to put ratios together instead of multiplying fractions; they must not speak of "the square root of two."

Mathematical modernity gets under way with Descartes' *Geometry*. By homogenizing what is studied, and by making the central activity the manipulative working of the mind,

rather than its visualizing of form and its insight into what informs the act of vision, Descartes transformed mathematics into a tool with which physics can master nature. He went public with his project in a cunning discourse about the method of well conducting one's reason and seeking the truth in the sciences; and this discourse introduced a collection of scientific try-outs of this method, the third and last of which was his *Geometry*.

For those who study Euclid and Apollonius in a world transformed by Descartes, many questions arise: What is the relation between the demonstration of theorems and the solving of problems? What separates the notions of *how-much* and *how-many*? Why try to overcome that separation by the notion of quantity as represented by a number-line? What is the difference between a mathematics of proportions, which arises to provide images for viewing being, and a mathematics of equations, which arises to provide tools for mastering nature? How does mathematics get transformed into what can be taken as a system of signs that refer to signs—as a symbolism which is meaningless until applied, when it becomes a source of immense power? What is mathematics, and why study it? What is learning, and what promotes it?

With minds that are shaped by the thinking of yesterday and of the days before it, we struggle to answer the questions of today, in a world transformed by the minds that did the thinking. We will proceed more thoughtfully in the days ahead if we have thought through that thinking for ourselves. Our scientific past is not *passé*.



The Husserlian Context of Klein's Mathematical Work

Burt C. Hopkins

I have to begin my remarks with the admission of my ignorance about their ultimate topic, which is neither Edmund Husserl's nor Jacob Klein's philosophy of mathematics nor, for that matter mathematics itself, but numbers. I do not know what numbers are. To be sure, I can say and read, usually with great accuracy, the numerals that indicate the prices of things, street addresses, what time it is, the totals on my pay stubs, the typically negative balance in my checkbook at the end of each month, and so on. Moreover, I know how to count and calculate with them, though usually with less accuracy no matter how much I try to concentrate on what I am doing.

But if I am asked or try to think about what they are when, for instance, I say my address is six hundred fifty-three Bell Street, or that my house has two bedrooms, or that I only have twenty bottles of wine left, it is obvious to me that I do not know what I am talking about. Of course, I know that my address refers to my abode, that my bedrooms are the rooms in the house with beds, and that my wine is something I drink solely for its medicinal purposes. But the six hundred fifty-three, the two, or the twenty, what are they? I do not know.

Likewise, if I am asked or try to think about what the numbers are with which I calculate, when for instance I think that because I need one-half pound of steak per person to feed each of my dinner guests, that I expect five guests, and that therefore I need two and a half pounds to feed my guests,

Burt Hopkins is Professor of Philosophy at Seattle University, and is Secretary of the International Circle of Husserl scholars. He is the author of a book, forthcoming from the University of Wisconsin Press, entitled *Edmund Husserl and Jacob Klein on the Origination of the Logic of Symbolic Mathematics: An Inquiry into the Historicity of Meaning*.

or when I think that the price of two and a half pounds of steak, at eight ninety-nine per pound is twenty-two dollars and forty-seven and a half cents, I know well enough what dinner guests, steak, and dollars and cents are. But the one half and the eight ninety-nine with which I calculate, and the two and a half and twenty-two and forty-seven and a half that are the results of my calculations, what are they? I do not know.

Some among us will say and, indeed have already said for a long time, that numbers, any numbers, are amounts of things, specifically, amounts of whatever it is that we count in order to answer the question how many of the things in question there are. In my examples above, to talk or think about two as the amount of my bedrooms, twenty as the amount of my bottles of wine, five as the amount of my dinner guests, eight ninety-nine as the amount of money it takes to buy a pound of steak, twenty-two and forty-seven and a half as the amount of money it takes to buy two and a half pounds of steak, certainly does seem to make sense. But what about my house address? Is six hundred fifty-three the answer to the question of how many of my house? Or is the one half pound of steak needed to feed each dinner guest the answer to the question how *many* of pounds of steak? The numbers here do not straight away appear to be telling us how many houses my house is, since the answer to that question is that it is not many at all but one; neither do numbers tell us how many pounds of steak are needed to feed each dinner guest because each does not need many pounds at all but only a half a pound. And, to complicate things deliberately, let us consider what it means, numerically speaking, if I have a negative balance in my checkbook at the end of the month. Whatever the amount of the negative number, let us hypothetically say it is an even negative one hundred six dollars, that certainly does not seem to provide the answer to the question of how many dollars and cents I have.

For the moment, however, let us push these concerns aside and assume that a number really is the amount of some-

thing. Moreover, let us suppose that the amount of something or its number first makes sense to us when we count more than one of something. Finally, let us suppose that any group, that is, any collection of more than one of something, no matter how big, has a number, which is to say, an exact or definite amount that answers the question how many with respect to the things in the group, and that this can be arrived at by counting. Does this really answer the question what a number is? Or, more precisely, does this really answer the question what numbers are? I say numbers because when we count we always use more than one number to arrive at the exact amount of something, even though once we arrive there we conclude the count by saying a single number.

What, then, are the numbers two, three, seven, or, if I have a lot of enumerative stamina, five hundred eighteen, that I say when, having nothing better to do, I count the grains of sand I have decided to put into separate piles? Or, what are the numbers two, three, seven, five hundred eighteen, that I say when I count the crystals of salt that I use to duplicate the numbers of grains of sand I previously put into piles? The things counted and therefore *their* amounts are not the same; that is to say, an amount of grains of sand and an amount of salt crystals are different things. But are their numbers the same? Is the two that is the amount of grains of sand the same two that is the amount of crystals of salt?

If number is supposed really to be the amount of something, and if the stress is placed on the *of something*, then number, as *its* (the something's) amount, would be nothing more or less than the *two grains of sand*, or, the *two crystals of salt*, both of which, being different somethings, would also be different numbers. This situation would be illustrated were I to say that there are a number of people I do not know in the room. If I proceeded to count them, whatever number I came up with would be not just an exact amount of anything whatever, but precisely the exact amount *of people* I do not know in the room.

On the other hand, if the stress were placed on the *how many* of the amounts in question, then two would indeed be how many grains of sand and crystals of salt there are in my smallest piles of each, and therefore their amounts and hence numbers would indeed be the same. This situation would be illustrated were I to say all dogs, if they are naturally formed, have the same number of legs, namely, four. The exact number is the same, though the legs are not, because, silly as it sounds to say this, different dogs have different legs.

Now some might want to put an end to this whole line of inquiry, but especially to my last question about whether numbers of different things are different or the same, by saying that the obvious answer is that what numbers are are abstract concepts. Hence they are really ideas, ideas that we can relate to different things when we want to count or want to apply the results of our calculations. But we do not have to. Thus I can add 4 to 6 and get 10. I can multiply 10 times 10 and get 100, and so on, without having to think or answer anybody's question about 4 of what, or 6 of what, or 10 of what, and so on. Just as it makes perfect sense to say that 2 plus 2 is 4, it does not make any sense to say 2 plus 3 is 4, because everybody who can count knows 2 plus 3 is 5.

Perhaps. But perhaps not. And this is where Husserl and then eventually Klein come in. But first Husserl. It is generally known that Edmund Husserl, the German philosopher who, as the founder of the so-called phenomenological movement in philosophy, was responsible for one of the two dominant approaches to philosophy in the last century (the other being so-called analytic philosophy), was originally a mathematician. Known likewise is that his first book, published in 1891, is titled *Philosophy of Arithmetic*.¹ However, the contents of this book are not so well known, because, among other reasons, soon after its publication both its author and Gottlob Frege had some very critical things to say about them.

For our purposes, however, the contents of the book are more important than their criticism of them. From beginning

to end, the book concerns the answer to the question whether numbers are really abstract concepts that make perfect sense without saying or thinking what they are numbers of; or whether to make sense *as numbers*, numbers have to be spoken or thought about as being *numbers of something*. Husserl began his answer to this question by first distinguishing between two kinds of numbers, one of which he called "authentic" and the other "symbolic." We will discuss the latter first, because even though, as its name suggests, it is the less authoritative kind of number, it is also the one with which we are usually more familiar. A number is symbolic in Husserl's understanding when the number and the sign used to designate it are indistinguishable. For example: 3. Most of us have no doubt been taught or learned that this *is* the number three rather than what it really is, which is a number sign or numeral. Indeed, even though most of us are also aware of other numerals, for instance, Roman numerals, my suspicion is that when we see such numerals we immediately interpret what they really mean in terms of our numeral system, a system that was actually invented by the Arabs. A symbolic number, then, is a number that most of us—with or without thinking about it—identify with the sign that we either write or read.

Most of us, that is, unless we have thought about the fact that the signs used to designate numerals, and therefore these numerals themselves, are based on what were originally and still remain conventions, even if they are no longer recognized as such. Different conventions mean different numerals, as we have just seen. But what about the numbers? Do different numerals mean different numbers? Many when faced with this question conclude no. Their thinking here is that numbers remain the same, however different the numerals that express them, because numbers are really concepts. Hence, no matter what numeral or word is used to express the number three, 'three' is a concept that remains the same because it is not identical with some numeral (which is subject to change) or with a word from some language (which is

subject to variability). Moreover, for just this reason number is a properly abstract concept: it remains identical with itself no matter what sign or word is used to express it.

Despite the sophistication of this view of number, it is dead wrong according to Husserl, because if numbers were really abstract concepts in this sense, then the most basic operation of arithmetic—addition—becomes unintelligible. For instance, if in adding the number two to the number two, what we are really adding is the abstract concept of the number two to another (!) abstract concept of the number two, then arriving at their putative sum, the abstract concept 'four', becomes a great mystery. Most obviously, there is the problem of how a concept that is supposed to remain identical can nevertheless change into another concept that is also supposed to remain identical. That is, the abstract concept 'two' is not supposed to be able to change as words and signs can and do, but to remain what it is, namely, the number two. Yet precisely this supposition has to be abandoned if talk of adding the concept of two to the concept of two to get the concept of four is to make sense, since when the number two is added to the number two the sum is not two number twos but the number four.

Indeed, it is precisely this consideration that led Husserl to the realization that authentic numbers are *not* abstract concepts. This is, admittedly, a difficult thought. What, then, are they, these authentic numbers, if they are not abstract concepts? Husserl's answer is that they are the definite amounts of definite things that have been grouped together by the mind. What kind of things? Literally any kind. What definite amounts? Pretty much the first ten, namely two, three, four, five, six, seven, eight, nine, ten. Why are zero and one not among these? Because they are not definite amounts, which in the case of zero is obvious, while in the case of one is less so but still fairly obvious, since one is not an amount; it is not many. What has the mind to do with authentic numbers? Plenty for Husserl, since not only do authentic numbers first show up in counting, but also, only those definite things can

be counted that have been grouped together by it to compose what Husserl, and as we shall also see, ancient Greek philosophers and mathematicians, called a multiplicity. Moreover—and this is the most important consideration for our purposes—an authentic number really is something that manifestly *cannot* be found in either the reality of the definite things that are counted or in any relationship among them.

This last point requires closer scrutiny. If we ask how it is that the first authentic number, two, is able to register a definite amount of definite things as 'two', in the sense that in saying or thinking the number two, the things in question are recognized as being exactly two with regard to their number, we are then asking about something that is different from the things that this number numbers as two. Husserl, following a long philosophical and mathematical tradition, refers to what is asked about in this question as the unity of this number. In asking it, we are asking how it is that distinct things, in this case two of them, can nevertheless be brought together as precisely this, namely, the two that is articulated by the number two. This all-important question about the unity of authentic numbers becomes more explicit when we consider Husserl's reason for thinking that only the first ten or so definite amounts can be authentic numbers. Husserl's reason is deceptively simple: the mind can only apprehend—*all at once*—each of the definite items that are numbered in a number when these things do not exceed ten. When the amount of definite things in a multiplicity of things exceeds ten, each one of them cannot be apprehended all at once by the mind as when, for instance, it counts thirty of them.

These considerations, do not provide us yet with Husserl's answer to the question of the unity of authentic numbers, but only address what is at stake in it. What is at stake is that in registering the definite amount of definite things, an authentic number is bringing them together in a way that cannot be explained by each of the things so brought together, no matter whether these are considered by themselves or as each relates to the other things. Each considered

by itself cannot explain it for the simple reason that authentic numbers are always amounts of something that is more than one. Considering the relations among things cannot explain it either—they may be side by side, or on top of, or bigger and smaller than one another, and so on, because none of these relations is even remotely numerical. Each definite thing, as one thing, can only be registered as having a number by being brought together with other definite things, each of which is also one, a bringing together which does not apprehend the things brought together singly but precisely as all together.

Husserl's explanation of how the mind does this is as simple as it is remarkable. The mind combines things into groups or collections in such a way that what is grouped or collected forms a whole that is different from the group members or collected items, even though as the whole of *just these* members or items, it is clearly related to them. For example, in a row of trees, a gaggle of geese, or a flock of birds, what is named by the row, gaggle, and flock is the whole Husserl has in mind, a whole that cannot be separated from what it is a whole of any more than it can be identified totally with it. Authentic numbers for Husserl are also comprised of wholes like this, so that the whole of the number three is clearly related to each of the things it registers as three, without, however, its being totally identical with them. Two distinct but related things are involved for Husserl here. One is the fact that there are groups and collections of this kind, and the other is that they have their origin in the mind. The talk about the mind's involvement in originating groups and collections of things does not mean that these things are only figments of the mind. Husserl only mentions the mind to explain something very specific, the fact that the unity, the whole of authentic numbers that we have been talking about, is neither an abstract concept nor something that can be explained by the definite things it registers as to their exact number. Once this is recognized, and *only* once it is recognized, are we then in a position to understand why Husserl

tried to explain this unity—this being a whole—of authentic numbers by the way the mind organizes and grasps things as groups and collections.

Husserl is very specific about this. The mind considers each thing that it groups or collects as something that belongs to the whole of whatever it is grouping and collecting. In the case of authentic numbers, it considers each as something, a certain one, without attending in the slightest to any other qualities that belong to what it collects. This is the case because unlike other groups and collections, which are groups of something specific, for example, geese or trees, authentic numbers are wholes of quite literally anything whatever. The moons of Jupiter, Homer's psyche, the city of Annapolis, etc., can be collected together and the collection registered as an authentic number—so long, of course, as the amount in the collection does not get too big. The process of forming authentic numbers, as well as other kinds of groups and collections, is expressed in language according to Husserl by the word "and," although both this process and the wholes it generates are for him most decidedly not anything linguistic. Hence the formation of a group, such as the group or whole of students in a room, comes about when one student and one student and one student and one student, and so on, are collected by the mind. Note well, however, that in this example not just anything can be grouped together, but only what belongs to the whole that is being grouped, namely, students. Likewise, the formation of a collection, the whole of which is the red objects in a room, comes about when one of any kind of red object and one of any kind of red object and one of any kind of red object, and so on, are collected by the mind. Again, as in the whole that is a group, not just anything can be collected. Finally, the formation of a collection, the whole of which is its number, comes about when something and something are collected, and then the process is stopped. More precisely, when the process of collecting is stopped after something and something are collected, the first authentic number, two, is the result. Likewise, when something and

something, and something are collected, the authentic number three is the result, and so on, until the limit of authentic numbers, ten, comes about.

Authentic numbers, then, are just such collections of something, which is to say, of anything that can be considered as simply *one* thing, without any other qualities or determinations of it being relevant to its belonging to a collection that is numbered. Since its being just one is the only thing relevant here, Husserl follows a long tradition and refers to these collected ones as “units.” Because authentic numbers are amounts of units, Husserl can explain on their basis what the understanding of numbers as abstract concepts cannot, namely, addition. Adding two and two involves the combination of ‘one unit and one unit’ with ‘one unit and one unit’, and hence, there is no mystery here, since ‘one unit and one unit, added to ‘one unit and one unit’, yields ‘one unit and one unit, and one unit, and one unit’, which is the authentic number four.

Before considering what Husserl thinks happens to numbers when their number exceeds ten and they are no longer authentic, let us step back for a minute. I have been talking now for some time about something I have acknowledged my ignorance of, namely, what numbers are. After considering two common views of them—one that considers them to be amounts of something and the other that considers them to be abstract concepts—I have turned our attention to the contents of Husserl’s book on the philosophy of arithmetic. The very question it attempts to answer is which of these two views of number is correct. So far I have pointed out that Husserl begins this investigation by distinguishing between authentic and symbolic numbers, both of which we have now discussed in some detail. Indeed, at this point it might seem that Husserl’s answer to the question of whether numbers are amounts of something or abstract concepts is pretty obvious, since authentic numbers can explain what abstract conceptual numbers cannot, namely, the basic operations of arithmetic. However, things are not that simple because when we con-

sider numbers greater than ten, Husserl thinks that they become inauthentic, as they cannot be authentic. They cannot be authentic, it will be recalled, because the mind cannot grasp more than ten things all at once. Symbolic numbers on Husserl’s understanding are therefore also in this sense inauthentic.

Even though they are inauthentic, however, symbolic numbers are not inferior, mathematically speaking, to authentic ones for Husserl. On the contrary, because they deal with numbers larger than ten, they come in handy any time calculation with large numbers is required, since without them, we would be reduced to counting units when we calculate with such numbers, which no doubt would be both tedious and time consuming. At the time Husserl wrote his first book, mathematicians and philosophers wanted an explanation how it was possible to do what no one denies can be done: to calculate with inauthentic numbers, which are symbolic and so in some sense are abstract concepts. In the first ten chapters of *The Philosophy of Arithmetic*, Husserl attempted to prove that authentic numbers and symbolic numbers are logically equivalent because each refers to the same objects, specifically, the collections of more than one unit that authentic numbers register the first ten amounts of. He argued that authentic numbers do this directly and symbolic numbers indirectly. When Husserl reached chapter eleven, however, he realized something that shook him to his depths, quite literally. (He later recounted a decade-long depression that ensued as a result.) He realized not only that symbolic numbers did not refer to the same objects as authentic ones—to collections of units—but also that the basic operations with quantities that are known, what he called general arithmetic, could only be explained on the basis of the very opposite of what he had argued in the first ten chapters. General arithmetic only makes sense if the numbers it uses are symbolic in the sense I discussed above, wherein the sign and the numerical concept are identical so that a number is interpreted to be the sign that we write or read—what Husserl called “sense-

perceptible” signs. Husserl also realized he had no idea of how this is possible. As he described it some two decades later, “how symbolic thinking is ‘possible’, how...mathematical...relations constitute themselves in the mind...and can be objectively valid, all this remained mysterious.”²

Thus we can say that, while recognizing that some numbers are clearly definite amounts of units and others are abstract, symbolic concepts because they do not refer to such units, Husserl came to see that he did not know what either of them really is. Husserl eventually thought he could solve this mystery by explaining the manipulation of symbolic numbers, or, more precisely, number symbols, as well as all mathematical symbols, in terms of what he called “the rules of a game,” rules that were invented not by mathematics but by logic. Mathematics thus came to be understood by Husserl as a branch of logic. Moreover, Husserl thought that all the rules invented by logic have their foundation in concepts that are true of other concepts, these other concepts, in turn, being true of anything whatever, that is, anything that can be experienced and therefore thought of as an individual object. As a consequence, Husserl’s eventual explanation of how symbolic mathematics is possible, in *The Crisis of European Sciences*, was really not so far from his failed first attempt at explanation. To be sure, his later explanation does not characterize number symbols as referring to the same objects that authentic numbers do, namely to units, but it did trace the truth of the logic that invented the rules for manipulating them to a basis in individual objects. This, we shall soon see, is a problem if we follow Jacob Klein’s mathematical investigations, which I am going to suggest can best be followed by first considering their Husserlian context.

The bulk of Klein’s mathematical investigations are contained in his *Greek Mathematical Thought and the Origin of Algebra*, which as many of you know was originally published in German as two long articles in 1934 and 1936 and translated into English in 1968 by St. John’s tutor and former Dean of the College, Eva Brann.³ The historical nature of the

topic announced by its title might initially seem to be far removed from what we now know is the systematic nature of Husserl’s topic in *Philosophy of Arithmetic*. However, one does not need to read very far in Klein’s book to discover that he understands the key to the historical investigation of his topic to lie precisely in the distinction between symbolic and non-symbolic numbers. Specifically, he expresses the view from the start not only that the symbolic number concept is something that makes modern, algebraic mathematics possible, but also, that symbolic numbers were entirely unknown to ancient Greek mathematicians and philosophers. The very point of departure for Klein’s investigation of the origin of algebra is therefore informed by his view that unless the non-equivalence of ancient Greek numbers and modern symbolic numbers is recognized, the change in the nature of the very *concept* of number that took place with the transformation of classical mathematics into modern mathematics in the sixteenth century will go unrecognized.

Klein’s thought here can be made clearer by closely considering precisely how he characterizes the difference between the ancient Greek numbers and the modern symbolic ones. Ancient Greek numbers or *arithmoi* are manifestly not abstract concepts, but rather *beings* that determine definite amounts of definite things. In contrast, symbolic numbers are characterized by him to be abstract concepts that do not refer to anything definite *except* that which is referred to by their sense-perceptible signs. When we consider the fact that Husserl articulated the difference between authentic and symbolic numbers in precisely these terms, the resemblance between Klein’s and Husserl’s view of the distinction between symbolic and non-symbolic numbers is striking. So striking is it in fact that some have drawn the conclusion that Husserl should be given precedence in this matter, as either the source or the major influence on Klein’s formulation of the distinction in question.⁴ These matters, however, are not so simple.

To begin, it is important to keep in mind that Husserl sought in *Philosophy of Arithmetic* to demonstrate the logical

equivalence of non-symbolic and symbolic numbers, whereas Klein's book *begins* with the insight that this is impossible, an insight, as we have seen, Husserl arrived at only reluctantly. Moreover, subsequent to his first book Husserl explains the relationship between symbolic and authentic numbers as a matter of logic. In his *Logical Investigations* (1900) and *Formal and Transcendental Logic* (1927), the relationship is explained as a matter of the translation of logical truths into rules for the manipulation of symbols. Logical truths for Husserl are rooted in concepts that are true of other concepts, other concepts that, in turn, can be traced back to truths that are rooted in individual objects. Husserl calls the rules established on the basis of this logic the "rules of a game," because even though they permit calculational operations on symbols that yield mathematically correct results, these operations and hence the very process of symbolic calculation have nothing to do with *insight* into *concepts* that pertain to the objects to which they are ultimately related, and therefore nothing to do with real knowledge. In a word, Husserl thought he resolved the issue of symbolic and authentic numbers by exposing symbolic calculation to be a "technique" whose cognitive justification can only be provided on the basis of the conceptual knowledge of individual objects that logic provides.

Klein, however, thought otherwise. He thought that it is impossible to explain the sharp distinction between non-symbolic and symbolic numbers on the basis of the knowledge—logical or any other kind—of individual objects. He conducted an historical investigation into how an aspect proper to ancient Greek *arithmoi* was transformed into modern symbolical numbers and concluded that the objects referred to by each kind of number are fundamentally different. The objects referred to by *arithmoi* are definite, which is to say, *individual*. The objects referred to by symbolic numbers are indefinite. They refer neither to individuals nor to their qualities and are therefore indeterminate. Individual objects, both Husserl and Klein agree, are objects that we encounter in our

experience of the world and thus, they can be pointed to. Moreover, their qualities, which either some (general qualities) or all (universal qualities) individual objects share, are qualities that, even though they are not individual, can nevertheless be spoken about in connection with the individual objects that we do encounter in the world. For instance, we can point to dogs and cats, each one of which is therefore individual. We can also talk about their general or universal qualities and relate these to the individual dogs and cats that we encounter in the world. Indeterminate objects, on the contrary, can never be encountered in our experience of the world, and therefore they cannot be pointed to. Husserl and Klein also agree that because their objects are not determinate in this very precise sense, symbolic numbers cannot have a direct relationship to any individual objects in the world or to their qualities. Finally, both Husserl and Klein agree that *symbolic numbers themselves, and not their indeterminate objects, are nevertheless encountered in the world*: namely, they are encountered as the sense-perceptible signs that, we mentioned earlier, many of us interpret *as* numbers.

Where Husserl and Klein disagree, or more accurately, where Klein *would* have had to express his departure from Husserl's understanding of the relationship between symbolic and non-symbolic numbers, *had* he chosen to do so, has to do with the possibility of theoretically clarifying the philosophical meaning of symbolic numbers. Husserl thought this could be done—indeed, he thought he did it in his two books on logic. Klein did not think it could be done. In fact, as we shall see, Klein's mathematics book explains why the very attempt to clarify theoretically the philosophical meaning of symbolic numbers and mathematical symbolism generally is doomed to failure. It is so doomed because all the concepts available to provide such a theoretical clarification, without exception, only make sense when the philosophers or anyone else using them are talking about individual objects and their qualities.

Klein explains why this is the case by establishing a fundamental difference in what he calls the "conceptuality"⁵ of

the concepts that belong to ancient Greek and modern science. He uses this term to articulate both the *way* in which the concepts proper to these respective sciences are structured and the status of their relationship to the non-conceptual realities of both the mind and the world. Klein thinks that despite the continuity discernable in the technical vocabulary of ancient Greek mathematics and philosophy and their modern counterparts, the mathematical and philosophical significance of each of the words in it is nevertheless *radically* divergent for the ancients and moderns. Thus in his view the fundamental significance of words like knowledge, truth, concept, form, matter, nature, energy, number, and so on is completely different for the ancient Greeks and the moderns because of the differences in the respective conceptualities of each. In other words, Klein thinks that these conceptualities shape the meaning of words, rather than the other way around, that is, rather than the significance of words shaping the structure of the conceptualities.

Klein locates the key example of the shift from the ancient Greek to the modern conceptuality in the transformation the concept of number undergoes in the sixteenth century. Prior to Viète's invention of the mathematical symbol, the concept of number according to Klein always only meant a definite amount of definite things, a meaning that was established by the ancient Greeks and that remained operative in both European mathematics and the Europeans' everyday praxis of counting and calculation until the invention of algebra. Klein claimed in his book—but did not elaborate—that this transformation is paradigmatic for the conceptuality that structures the modern consciousness of the world.

Before considering in more detail Klein's account of this exemplary transformation of the conceptuality of number, a few words about the potentially misleading talk of the "concept" of number are in order. Klein engages in such talk when his investigation is *comparing* what he refers to as the different "number concepts" of ancient Greek and modern mathe-

matics. It is potentially misleading because for Klein the most salient difference between in these number concepts is located in the fact that the ancient Greek "concept" of number is precisely something that is *not* at all a concept, but a *being*, while the modern "concept" of number is precisely something that is not a being but a *concept*. The object of the word "concept" in these comparative contexts is, I think, clearly the "conceptuality" of ancient Greek and modern numbers, which means it would be a mistake to attribute to Klein in such contexts the thought that in ancient Greek and modern mathematics numbers are concepts, albeit different in kind. Klein, however, also talks about the ancient Greek "*arithmos*-concept" (*arithmos-Begriff* or *Anzahl-Begriff*) and the modern "number-concept" (*Zahl-Begriff*), which again is potentially misleading, for the same reasons. Yet here, too, careful consideration again discloses that such talk always occurs within the context of his comparison of what he presents as the different ancient Greek and modern *characterizations* of numbers, *only one of which formulated them as concepts*.

Before elaborating Klein's account of these different characterizations, I want to raise and then answer one more question. From what perspective was Klein able to compare the ancient Greek and modern numbers? Klein, after all, was neither ancient nor Greek but, by his own admission, thoroughly modern. How, then, was he able nevertheless to get sufficient distance from the presuppositions that inform his modernity, from his modern outlook and consciousness, such that he could encounter and investigate what, again by his own admission, are the radically different presuppositions of the ancient Greeks?

I think the answer to this question can be found in the single reference in Klein's published work to Husserl's *Philosophy of Arithmetic*.⁶ It occurs in "Phenomenology and the History of Science," an article Klein wrote for a memorial volume of essays on Husserl's phenomenology published in 1940, two years after Husserl's death. In this article Klein did

not at all hesitate to articulate the philosophical significance of his 1934 and 1936 investigations of ancient Greek mathematics and the origin of modern algebra in terms of Husserl's last writings, which were published in 1936 and 1939.⁷ In these writings Husserl traces the cause of the crisis of European sciences to their failure to grasp properly the scope and limits of scientific methods that are rooted in modern mathematics for understanding the non-physical, which is to say human spiritual reality together with the world of its immediate concerns. I will come back to this theme at the end of my remarks. I want to focus for now on Klein's reference to what he characterizes in this article as Husserl's "earliest philosophical problem," namely "the 'logic' of symbolic mathematics." He asserts, "The paramount importance of this problem can be easily grasped, if we think of the role that symbolic mathematics has played in the development of modern science since the end of the sixteenth century." Klein concludes his remarks on "Husserl's logical researches" by saying that these researches "amount in fact to a reproduction and precise understanding of the 'formalization' which took place in mathematics (and philosophy) ever since Viète and Descartes paved the way for modern science."

Klein's qualification that Husserl's logical researches "amount . . . to" both a "reproduction" and "precise understanding" of the "formalization" in mathematics initiated by Viète contains the key to my answer to the question. It indicates that Klein, but *not* Husserl, was aware of the historical significance of these researches. The result of the "formalization" in mathematics referred to here by Klein concerns the "indeterminacy" of the object of mathematical symbols that I called attention to earlier, the absence of any direct reference to both individual objects and their general and universal qualities that is the mathematical symbols' most salient characteristic. Because Klein thought Husserl's logical researches in *Philosophy of Arithmetic* amount in fact to the reproduction and precise understanding of the historical genesis of this formalization, it would follow from this that Klein under-

stood that work's investigation of the relationship between authentic and symbolic numbers to mirror what his own research presents as the relationship between the ancient Greek *arithmos* and the modern symbolic number.

This being the case, the question of image and original suggested by my mirror metaphor arises, namely, did Husserl's investigations mirror Klein's or Klein's Husserl's? Here I think chronology is relevant, which would suggest that what enabled Klein to encounter the presuppositions of both his own modern as well as the ancient Greek conceptuality was Husserl's reluctant discovery in *Philosophy of Arithmetic* that the formal conceptual status of symbolic numbers cannot be rendered intelligible in terms of authentic numbers. By saying this, however, I want to emphasize in the strongest terms possible that I am *not* suggesting what some others have suggested, namely, that what makes Klein's comparison of the conceptuality of the ancient Greek and modern numbers possible is his *projection* of Husserl's "concepts" of authentic and symbolic numbers back into the history of mathematics.⁸ On the contrary, I want to suggest and then develop a much more subtle and more radical claim. Husserl's failure provided Klein with the guiding clue that enabled him to trace and illuminate certain historical dimensions of that very failure. Klein detected an historical transformation of non-symbolic numbers: an *aspect* of their conceptuality was transformed into symbolic numbers. This discovery resulted in a more definitive philosophical account of both kinds of numbers. Moreover, Klein discovered something of which Husserl had not the slightest inkling, namely, that the formal conceptuality of symbolic numbers, and symbolic conceptuality in general, cannot be made intelligible on the basis of theoretical concepts traceable to ancient Greek science. And that discovery highlighted something more ominous: built into symbolic cognition is the misguided self-understanding that evaluates its own cognitive status as the ever-increasing perfection of ancient Greek science's theoretical aspirations.

Klein's account of the origination of algebra is really the story of the simultaneous invention of the mathematical symbol, the symbolic formulae it makes possible, and the resultant novel art of symbolic calculation. It is a story with a such a complex array of intricate subplots and such a diverse cast of characters that it is not always easy to follow the action, especially given its approximately two-thousand-year time span. It is also a difficult story to follow, since not only are there really no good and bad guys to identify with or to dislike, but its beginning as well as its ending is obscure. The action takes place, for the most part, in the realm of pure beings and pure concepts, a realm that is invisible to the eyes and in which all the actors are likewise invisible and therefore, with some justification, referred to by many as "abstract." Despite its otherworldly aura, it is a story well worth trying to follow because what it is about is the origin of the *mistaken identity* of the very technique that has enabled mathematical physics and the technology spawned from it quite literally to transform the world. Unraveling the plot is an exercise in discovering the true identity of this technique and, in the process, rediscovering something essential about our relation to the world that the events surrounding the technique's origination continue to make it easy for us to forget.

Turning now to the story: we have already seen that non-symbolic and symbolic numbers are key players in Klein's tale. Guided by our discussion of Husserl's account of them, we are in a position to see what it means to say with Klein that the non-symbolic numbers of the ancient Greeks are not concepts, let alone abstract concepts: being definite amounts of definite things, *arithmoi* were initially understood not only to be inseparable from things but also to be what is responsible for the well-ordered arrangement proper to all their parts and qualities. In other words, they were understood by the Pythagoreans as the very being of everything that is. To be is to be countable, and because to be countable each thing has to be one, the one was very important, as was the odd and the

even, since whatever is counted ends up being odd or even. The one or the unit (*monas*), as something without which counting is impossible, is therefore the most basic principle (*archê*) of *arithmos*. The odd and the even, which order the *arithmos* of everything countable, insofar as it *has* to be one or the other, manifest the first two *kinds* (*eidê*) of *arithmoi*. Moreover, since the even can be divided without ever arriving at a final *arithmos*, while the odd cannot be divided evenly at all, because a one is always left over, these two kinds are understood, respectively, as unlimited and limit.

It is important to note here three things, according to Klein: (1) the *arithmoi* are inseparable from that which is countable; (2) the most basic principle as well as the kinds of *arithmoi* are not themselves *arithmoi*. In other words, they are not numerical if by numerical we understand, as the ancient Greeks did, number to be a definite amount of definite things; and (3) the *arithmoi*, being inseparable from what is countable, are not abstract entities, and because they are different from both their most basic principle and their kinds, they are not even remotely "conceptual," assuming for the moment that it is even appropriate to use this adjective to refer to the *archê* and *eidê* of *arithmoi*. It is important to note these three things because, on Klein's telling, no matter how much the ancient Greek characterization of the mode of being of *arithmoi* changes in what become, in Plato and Aristotle, the two paradigmatic ways of its understanding it, these three things about the *arithmoi* remain constant. In Klein's words, "All these characterizations stem from one and the same original intuition [*Anschauung*], one oriented to the phenomenon of counting."⁹

While the discovery of incommensurable magnitudes brought to an end the Pythagorean dream of a world in which being counted was identical with being measured, Plato's positing of the invisible, indivisible, and therefore sensibly pure mode of being of the *archê* of *arithmoi* brought into being another dream, the dream of what Klein calls an *arithmological* ordering of the *eidê* responsible for *arithmoi* as

well as anything else that is a being. As Klein relates it, Plato realized that the Pythagorean account of the *archê* of *arithmoi*, the one or the unit, as something that is quite literally inseparable from the sensible things that are countable on its basis, presented an obstacle to understanding the true relation of *arithmoi* to the soul when it counts. This is the case because Plato must have noticed that even *before*, the soul counts each one of the definite things that it perceives, it already has some understanding of the *arithmoi* it employs in arriving at the *arithmos* of such things. According to Klein, Plato must have thought this possible because *prior* to counting sensible things the soul has available to it *arithmoi* that are definite amounts of *intelligible (noêta)* units, intelligible in the sense that they cannot be seen with the eyes, cannot be divided like the things seen with the eyes, and cannot be unequal like the things seen by the eyes. Just like the Pythagoreans' sensible *arithmoi*, these intelligible *arithmoi* are either odd or even, though unlike the Pythagorean *eidê*, those belonging to intelligible *arithmoi* are likewise intelligible, and thus cannot be seen with the eyes.

Now it has to be stressed here that absolutely *nothing* is either abstract or general about Plato's intelligible *arithmoi*. They are not abstract because they are not lifted off anything. They are not general because they are precisely definite amounts of definite things, albeit in this case the things are *noeta*. Moreover, they are not general because, just like the Pythagorean *arithmoi*, they are *not* concepts: each *arithmos* is a definite whole, the unity of which is exactly so and so many intelligible units. What allows intelligible numbers to be used in the counting of anything whatever is what Plato's Socrates never tired of pointing out to his interlocutors, namely, that the true referents of our speech, in counting off amounts of things or in anything else, are not sensible but intelligible beings. Thus in counting what the soul is really aiming at when it counts off in speech, *the definite amounts of definite sensible things, things that in being counted are treated as sensible units, are definite amounts of intelligible units*. Indeed, it

is precisely this state of affairs that allows the soul to count anything that happens to be before it, since the true units of its counting are not those that can be seen but precisely those that can only be thought.

Plato's way of explaining how the availability of intelligible *arithmoi* to the soul enables it to count anything whatever means that these intelligible units are manifestly *unlike* the units in Husserl's authentic numbers. Plato's intelligible units explain the ability of *arithmoi* to count anything whatever because they, and not the "whatever," are each *arithmos*' true referent. In other words, for Plato—and for that matter, for all the ancient Greeks *including* Aristotle¹⁰—there is no such concept of *any thing, or any object, or any being whatever*. Such a concept, as we have seen, is explained by Husserl in terms of an abstracting activity of the mind that is so powerful it is powerful enough to create a concept so general that literally anything whatever (*Etwas überhaupt*) can "fall under it." This is to say, for Husserl the mind is able to create a *formal* concept that has absolutely *no determinate* reference to any individual thing in the world or to the general and universal qualities of such things. Klein's point, and in my judgment the point is the fulcrum upon which the story told in his math book pivots, is that until Viète invented algebra, the power behind this abstraction—what Klein calls a symbol generating abstraction—was something that the world had never seen before.

Husserl's concept of the units in authentic numbers is therefore modern, which is something Klein was not only no doubt aware of, but it is also no doubt the reason why Klein silently passed over in silence Husserl's investigations of the logic of symbolic mathematics. Indeed, Klein's book also bypassed any reference to Husserl's concept of intentionality when he articulated the difference between the conceptuality of non-symbolic and symbolic numbers in terms of the medieval concept of intentionality, and, again, no doubt it was for the same reason: a part of the composition of Husserl's concept of intentionality was already determined by the very for-

mality that Klein was investigating the origin of, in the shift from non-symbolic to symbolic numbers. Once this is realized, Klein's use of the medieval concepts of first and second intentional objects, rather than Husserl's concepts of straightforward and categorial intentional objects, to talk about the difference in the mode of being of non-symbolic and symbolic numbers makes perfect sense.

The transformation of an *aspect* of the ancient Greek *arithmoi* into the modern, symbolic numbers, however, does not make perfect sense. That is, Klein's account of the shift in the referent of ancient Greek and modern symbolic numbers is something that does *not*, and indeed *cannot*, render theoretically transparent the philosophical meaning of either the shift or the different numbers in question. The characterization of ancient Greek *arithmoi* in terms of their direct encounter with either sense perceptible objects encountered in the world or with intelligible objects encountered in the soul—what Klein reports the medievals called first-intentional objects—does not explain *what* such *arithmoi* are, in the precise sense of how it is that the different *arithmoi* render *intelligible* the different definite amount that characterizes each *arithmos*. Indeed, Klein never suggests that the concept of a first-intentional object can do this. Likewise, the characterization of symbolic numbers as having their referent in the mind's conception, a conception the medievals called the object of a second intention, does not render theoretically perspicuous what symbolic numbers are, either. To characterize symbolic numbers as pertaining to that *aspect* of *arithmoi* that concerns the “how many” of something, while at the same time *no longer* pertaining to its exact determination that each *arithmos* brings about, does not clarify theoretically what a symbolic number is. In other words, pointing out that the conceptuality of symbolic numbers disregards both the units of the something whose definite amount it is the province of *arithmoi* to determine and the exact amount of these units that each *arithmos* registers, does not explain what a symbolic number is—and neither does Klein's account

of how more than this shift is involved in its conceptuality. Klein affirms that what is required in order for us to be dealing with a symbolic number is that the second-intentional mode of being of the “how many,” which brings into the world for the first time a *formal* concept because it is now shorn of any connection to either first-intentional objects or the exact determination of their amount, *be expressed in a sense-perceptible sign that is grasped by the mind as the object of a first intention*.

What Klein's talk of objects of first and second intentions accomplishes is to call attention to something that nobody else in the twentieth century had seen, namely, that the invention of symbolic cognition represents nothing less than a reversal of the pre-modern relationship between concepts and objects: what were concepts for the ancient Greeks are now objects and what were objects for them are now concepts. Modern “theoretical thinking,” being symbolical, is thus necessarily blind to this reversal. Ancient “theoretical thinking,” not being symbolical, is likewise necessarily blind to it. It does not follow from this, however, that thinking *per se* must remain blind to it. On the contrary, for a thinking that is on its guard against falling victim to the most shameful ignorance, that is, to thinking it knows what, in truth, no mortal can know, not only is the shift in conceptuality articulated by Klein something that can be seen, but once seen, it is something that the soul's *phronesis* can never forget.

If we had more time, I would continue my remarks by calling attention to what I think is the key to Klein's account of how something like a symbol generating abstraction was able to come into the world, namely on the basis of Viète's, Stevin's, Descartes's, and Wallis's formulating the method of an art that permits calculation with what the ancient Greeks characterized not as *arithmoi* but as their *eidê*. And, indeed, I would call attention to the fact that, for Klein, with this not only do the true objects of mathematics become conceptual, that is, formal, but also, such concepts at the same time *become numerical*. Finally, I would call attention to the par-

allel Klein draws between the breaking of the bounds of the intelligibility proper to the *logos* that is the result of Plato's formulation of the *eidê* of beings as having an *arithmological* structure and the similar breaking of these bounds by Viète's numerical formulation of the symbolic calculation with the *species* of numbers in symbolic cognition,¹¹ and then ask the following question: Do both of these attempts to comprehend beings theoretically transcend the limits of what can be spoken of intelligibly for the same simple reason, namely, that their theories presuppose a knowledge of what no mortal can claim in truth really to know, namely that most wondrous gift of the gods to humans—one and number?

Before I conclude, I would like to return very briefly to Klein's memorial essay on Husserl that I mentioned earlier and to the matter of his articulation in that essay of the philosophical meaning of his mathematical investigations in terms of Husserl's of last writings, the so-called crisis texts. It is important to draw attention here to the chronology of Klein's mathematical investigations and Husserl's last writings, because it is only in these writings that Husserl connects the two themes that had already informed Klein's *earlier* investigations of the history of mathematical concepts. Prior to 1936, when the second part of Klein's investigations were published, Husserl therefore had not yet recognized what Klein had already recognized, and indeed investigated extensively, guided as I have suggested by Husserl's first—and failed—investigation of the relationship between non-symbolic and symbolic numbers. Klein recognized the connection between the philosophical meaning of mathematical concepts and the history of their origination.

Notes

¹ Edmund Husserl, *Philosophie der Arithmetik*, ed. Lothar Eley, Husserliana XII (The Hague: Nijhoff, 1970), 245; English translation: *The Philosophy of Arithmetic*, trans. Dallas Willard (Dordrecht: Kluwer, 2003).

² Edmund Husserl, *Introductions to the Logical Investigations*, ed. Eugen Fink, trans. Philip J. Bossert and Curtis H. Peters (The Hague: Martinus Nijhoff, 1975), 35. German text, "Entwurf einer 'Vorrede' zu den 'Logischen Untersuchungen' (1913)," *Tijdschrift voor Philosophie* (1939): 106-133, here 127.

³ Jacob Klein, *Greek Mathematical Thought and the Origin of Algebra*, trans. Eva Brann (Cambridge, Mass.: M.I.T. Press, 1969; reprint: New York: Dover, 1992). This work was originally published in German as "Die griechische Logistik und die Entstehung der Algebra" in *Quellen und Studien zur Geschichte der Mathematik, Astronomie und Physik*, Abteilung B: *Studien*, vol. 3, no. 1 (Berlin, 1934), pp. 18–105 (Part I); no. 2 (1936), pp. 122–235 (Part II). Hereinafter referred to as GMTOA.

⁴ See the following: Hiram Caton, who claims that "Klein projects Husserl back upon Viète and Descartes," (*Studi International Di Filosofia*, Vol 3 (Autumn, 1971): 222-226, here 225; J. Phillip Miller, who writes "Although Husserl's own analyses move on the level of a priori possibility, Klein's work shows how fruitful these analyses can be when the categories they generate are used in studying the actual history of mathematical thought," (*Numbers in Presence and Absence* [The Hague: Martinus Nijhoff, 1982], 132; Joshua Kates' account is more circumspect, as he notes "[i]t is difficult to capture adequately . . . how much of Klein's understanding of Greek number is already to be found in Husserl, despite the important differences between them," ("Philosophy First, Last, and Counting: Edmund Husserl, Jacob Klein, and Plato's Arithmological *Eidê*," (*Graduate Faculty Philosophy Journal*, Vol. 25, Number 1 (2004), 65-97, here 94.

⁵ This is the literal translation of the word in question here, "*Begrifflichkeit*," which is rendered for the most part as "intentionality" in the English translation GMTOA. Because of this, the point I make below about the "concept of number" and "number concepts" will be more familiar to readers of Klein's original German text.

⁶ Jacob Klein, "Phenomenology and the History of Science," in *Philosophical Essays in Memory of Edmund Husserl*, ed. Marvin Farber (Cambridge, Mass.: Harvard University Press, 1940), 143–163; reprinted in Jacob Klein, *Lectures and Essays*, ed. Robert

B. Williamson and Elliott Zuckerman (Annapolis, Md.: St. John's Press, 1985), 65–84, here 70.

⁷ Edmund Husserl, "The Origin of Geometry," in *The Crisis of European Sciences and Transcendental Phenomenology*, trans. David Carr (Evanston, Ill.: Northwestern University Press, 1970). The German text was originally published in a heavily edited form by Eugen Fink as "Die Frage nach dem Ursprung der Geometrie als intentional-historisches Problem," *Revue internationale de Philosophie I* (1939). Fink's typescript of Husserl's original, and significantly different, 1936 text (which is the text translated by Carr) was published as Beilage III in *Die Krisis der europäischen Wissenschaften und die transzendente Phänomenologie. Eine Einleitung in die phänomenologische Philosophie*, ed. Walter Biemel, Husserliana VI (The Hague: Nijhoff, ¹1954, ²1976). Edmund Husserl, "Die Krisis der europäischen Wissenschaften und die transzendente Phänomenologie. Eine Einleitung in die phänomenologische Philosophie," *Philosophia I* (1936), (the text of this article is reprinted as §§ 1–27 of the text edited by Biemel).

⁸ See note 4 above.

⁹ *GMOT*, 54.

¹⁰ Aristotle's dispute with Plato over the mode of being of the *arithmoi* studied by the discipline of mathematics was about the origin of the units that they are the definite amounts of, and not whether theoretical *arithmoi* are definite amounts of units.

¹¹ "As Plato had once tried to grasp the highest science 'arithmologically' and therewith exceeded the bounds set for the *logos* (cf. Part I, Section 7C), so here [in Viète's invention of symbolic calculation] the 'arithmetical' interpretation leads to . . . the conception of a *symbolic* mathematics," the implication being, of course, that such a conception exceeds the same bounds as did Plato's attempt to grasp dialectic in terms of the *arithmoi eidetikoi* (*GMOT*, 184).



Words, Diagrams, and Symbols: Greek and Modern Mathematics or "On the Need To Rewrite The History of Greek Mathematics" Revisited

Sabetai Unguru

Mademoiselle de Sommery, as Stendhal tells us in *De l'Amour*, was caught "en flagrant delit," i.e., *in flagranti*, by her lover, who was shaken seeing his "amante" bedding down another man. Mademoiselle was surprised by her lover's angry reaction and denied brazenly the event. When he protested, she cried out: "Oh, well, I can see that you no longer love me, you would rather trust your eyes than what I tell you."

The historian of mathematics should behave like Mademoiselle's lover: believe his eyes and not what mathematicians-turned-historians tell him about the texts he studies. There are optical illusions, it is true, but they are to be preferred to the illusionary mental constructs of the mathematical historians. A text is a text is a text. Moreover, metaphors aside, not everything is a text and it behooves the cultural critic, and surely the historian, to relate only to written records as texts. Furthermore, texts are about something definite and not every conceivable interpretation suits them all. The *Conica* is about conic sections not about women. It deals with three (or four) kinds of lines obtained by cutting a

Sabetai Unguru is Professor Emeritus and former Director of the Cohen Institute for the History and Philosophy of Science and Ideas at the University of Tel-Aviv. He is the author of a controversial study with the title "On the Need to Rewrite the History of Greek Mathematics," as well as of a two-volume introduction to the history of mathematics. He is (with Michael Fried) the co-author of *Apollonius of Perga's Conica: Text, Context, Subtext*.

conic surface with a plane, lines baptized by Apollonius as parabolas, ellipses, hyperbolas, and opposite sections. It does not deal with a classification of women according to their degree of perfection, as seen by a male chauvinist pig, hiding his true intentions behind the cloud of mathematical jargon.

Texts are made of words, sometimes accompanied by illustrations; in the case of Greek geometry, the texts are made of words and diagrams, *C'est tout*. And the diagram *is*, in a very definite sense, the proposition. The words accompanying it serve to show how the diagram is obtained; they provide us the diagram *in statu nascendi*. But in principle it would be possible to supply the absent words to an extant diagram, though I doubt the possibility of understanding a reasonably sophisticated geometrical text in the absence of diagrams. If there is a manipulative aspect to Greek geometry, and I think there is, it resides in the construction, the bringing into being of the diagram, while the steps of the process are supplied by the words accompanying the diagram, typically in the *kataskene* (construction).

There are no true symbols in a Greek mathematical text. What looks like symbols to the untrained modern eye are actually proper names for identifying mathematical objects. They are not symbols, and cannot be manipulated, as algebraic symbols are. Even in Diophantus, which is a late and special case, his so-called symbols are actually verbal abbreviations, making his *Arithmetica* an instance of syncopated, not symbolic, algebra, in Nesselmann's tripartite division (*Die Algebra der Griechen*).

In modern, post-Cartesian mathematical texts, on the other hand, there are words and diagrams and symbols, but the actual necessity of the first two ingredients is minimal, serving heuristic, pedagogical, and rhetorical needs that can be dispensed with, leaving the text in its symbolic nakedness. It is no exaggeration to see modern mathematics as symbolic, while ancient mathematics cannot be seen historically in symbolic terms. This being the case, interpretations of ancient mathematical texts relying on their symbolic transmogrifica-

tion are inadequate and distorting. They are ahistorical and anachronistic, making them unacceptable for an understanding of ancient mathematics in its own right.

An ancient text—mathematical, philosophical, literary, or whatever—is the product of a culture foreign to ours, whose concerns, values, aims, standards, ideals, etc., are, as a rule, as alien to ours as can be. Though, inescapably, all history is retrospective history, the approach described and decried by Detlef D. Spalt, in his *Vom Mythos der mathematischen Vernunft* (1981), as *Resultatismus*, or, “Orwellsche 1984-Geschichtsschreibung für den grossen Bruder Vernunft,” that takes its bearings and criteria from what it sees as the *modern* outcome of a lengthy, linear, and necessary evolution of concepts and operations, looking always back at the past in light of its modern offspring, is necessarily a highly distorting approach, since it adopts unashamedly the perspective of the present to (in this order) judge and understand the past. Taken at face value, Percy W. Bridgman's statement that the past has meaning only in terms of the present is simply not true. The historian's stance is rather the opposite: the present has meaning only in terms of the past. *Prima facie* and on principled grounds therefore, an interpretation of any written, reasonably extensive document belonging to an ancient culture that results in its totally unproblematic and absolute assimilation to our own is suspect. That this is so is, more or less, acceptable when it comes to cultural artifacts other than mathematical ones, which seem to enjoy the privilege of perdurability and universality. The immunity from cultural specificity that mathematical truths command stems from the prevailing view that their outward appearance—their packaging, as it were—and their purely mathematical content—the packaged merchandise, as it were—are neutral, unrelated, and mutually independent items. It is a calamitous view and the root of all evil in the historiography of mathematics.

But there is another entrance into an ancient text, mathematical or not, one that does no violence to it, that does not break the inviolable unity of form and content and then enter

victoriously through the shambles it created, with the claim that the text is now understood; it is rather an ingress that accepts willingly and respectfully the unity of the text without pulverizing its inseparable aspects, bringing to its understanding both critical acumen and full acceptance of its outward appearance. This is the historical approach. The mathematical and historical approaches are antagonistic. Whoever breaks and enters typically returns from his escapades with other spoils than the peaceful and courteous caller.

Let me be more specific. Faced with an ancient mathematical text, the modern interpreter has an initial choice. First is the mathematical approach. It consists of two steps: (1) try to find out how one would do it (solve the problem, prove the proposition, perform the construction, etc.) and then (2) attempt to understand the ancient procedure in light of the answer to step (1). Instead of this preeminently mathematical approach, however, the modern interpreter can refuse to decipher the text by appealing to modern methods, using for its understanding only ancient methods available to the text's author. This is the historical approach. Needless to repeat, the spoils of interpretation differ according to the two approaches followed. The longstanding traditional approach has been the mathematical, though in the last three decades or so the historical approach is gaining increasingly more and more ground and, at least in the domain of ancient Greek mathematics, seems to be now the prevailing one. What seems certain is that in practice no compromise is possible between the mathematical and historical methodological principles. Adopting one or the other has fateful consequences for one's research, effectively determining the nature of the results reached and the tenor of the inferences used in reaching them.

What I am saying, then, is that despite the numerous and varied styles of writing the history of mathematics throughout the centuries, it is possible to group all histories of mathematics into two broad categories, the "mathematical" and the "historical." The former sees mathematics as eternal, its

truths unchanging and unaffected by their formal appearance, and sees the mathematical kernel of those truths as being independent of their outward mode of expression; the latter denies this independence and looks upon past mathematics as an unbreakable unity between form and content, a unity, moreover, that enables one to grasp mathematics as a historical discipline, the truths of which are indelibly embedded in changing linguistic structures. It is only this approach that is apt to avoid anachronism in the study of the mathematics of other eras.

To make this a historical talk, what is needed are specific, historical examples, supporting the preceding generalizations. I have offered numerous such examples in my published work, most recently in the book I published with Michael Fried, *Apollonius of Perga's Conica: Text, Context, Subtext* (2001). However, I would like now to enrich my offerings by drawing illustrations from historical sources less drawn upon in the past. One such source is Euclid's *Data*.

In one of the attacks launched against a notorious article of 1975, "On the Need to Rewrite the History of Greek Mathematics," Hans Freudenthal argues that, had its author been aware of the existence of the *Data*, he "would never have claimed there were no equations in Greek geometry." For Freudenthal, and not only for him, the *Data* is a "text-book on solving equations." He summarizes the 94 propositions contained therein in a succinctly and strikingly epigrammatic statement: "Given certain magnitudes a , b , c and a relation $F(a, b, c, x)$, then x , too, is given." But the fact remains that Greek geometry contained no equations. One cannot find even one equation in the entire text of the *Data*. Proof (as the Hindu mathematician would say): "Look!" Unless one has at his disposal the algebraic language and the capacity to translate into it, it is impossible to sum up this little treatise of rather varied content as offhandedly as Freudenthal has done. Indeed, had Euclid at his disposal Freudenthal's functional notation, it is rather easy to infer

that he would not have needed 94 propositions to get his point across.

Each case in Euclid's *Data* is unique, having its own method of analysis, and none is subsumable under or reducible to other cases, though, of course, later propositions rely on earlier ones. Thus, "The ratio of given magnitudes to one another is given" (proposition 1) and "If a given magnitude have a given ratio to some other magnitude, the other is also given in magnitude" (proposition 2)—to use perhaps the simplest illustration possible—are not for Euclid both instances of "Given a, b, c and $y = F(a, b, c, x)$, x is also given," but are two different problems, interesting in their own right, having their own solutions. Of course, Freudenthal's description is mathematically correct. Historically, however, it is wanting. Heath is much more to the point when he says:

The *Data*...are still concerned with *elementary geometry* [my italics], though forming part of the introduction to higher analysis. Their form is that of propositions proving that, if certain things *in a figure* [my italics] are given (in magnitude, in species, etc.), something else is given. The subject-matter is much the same as that of the planimetric books of the *Elements*, to which the *Data* are often supplementary.

This is what the *Data* is, not a textbook on solving equations, but a treatise presenting another approach to elementary geometry—other than that of the *Elements*, that is.

As an example of this characterization of the *Data*, I shall present proposition 16, in the new translation of C. M. Taisbak [*Euclid's Data or The Importance of Being Given*, (Copenhagen, 2003)]:

If two magnitudes have a given ratio to one another, and from the one a given magnitude be subtracted, while to the other a given magnitude be added, the whole will be greater than in ratio to the remainder by a given magnitude (p. 75).

First we must clarify the meaning of the expression "greater than in ratio...by a given." Definition 11 of the *Data* reads:

A magnitude is by a given greater than in ratio to a magnitude if, when the given magnitude be subtracted, the remainder has a given ratio to the same. (p. 35)

The meaning of this definition is, according to Taisbak, as follows (p. 57): " M is by the given G greater than the magnitude L which has to N a given ratio;" in other words, $M = L + G$ and $L:N$ is a given ratio.

Back to the proof of prop. 16:

Let two magnitudes AB, CD have a given ratio to one another. From CD let the given magnitude CE be subtracted and to AB let the given magnitude ZA be added. Then, the whole ZB is greater than in ratio to the remainder DE by a given magnitude.

$$\begin{array}{cccc} Z & & A & H & B \\ \hline C & & & E & D \end{array}$$

Now, since the ratio $AB:CD$ is given and AH can be obtained, by Dt. 4, from $AH:CE::AB:CD$, i.e., $AH:CE$ is also given, it follows that CE is also given. Hence, AH is given (by Dt. 2). But AZ is given; therefore the whole ZH is given (by Dt. 3). Since $AH:CE::AB:CD$, the ratio $HB:ED$ of the remainders is also given, by V. 19 and Def. 2, i.e., $HB:ED::AB:CD$. But HZ is given; therefore, ZB is by a given greater than in ratio to ED (Def. 11), *Q.E.D.*

No algebra appears here, and although the language of givens, the idiosyncratic concept "by a given greater than in ratio" and the sui generis concatenation of inferences burden the understanding, the proposition is clear and rather simple. It is graspable as it stands, without any appeal to foreign

tools. And yet, Clemens Thaeer, in his *Die Data von Euklid* (1962), proceeds as follows to its clarification:

Let $a:b=k$, then the proposition claims that if $x=ky$, then $(x+c)=k(y-d)+(c+kd)$, which is, of course, correct mathematically, but this blatantly algebraic procedure betrays the *Data*. It is a betrayal since ratios in Greek mathematics are not real numbers, i.e., the initial substitution $a:b=k$ is not kosher. Though there is some controversy about the status of ratios in the *Elements*, with respect to their being two, or four-place relations, their status in the *Data* is uncontroversial: a ratio P:Q is an individual item “however impalpable. A and B are magnitudes, most often (and least problematically) understood to be line segments; one may think of them as *positive real numbers*, that is as *lengths* of line segments, while remembering that the Greek geometers could not think like that, for want of such numbers” (Taisbak, *Euclid's Dedomena*, p. 32). Taisbak claims that distorting “clarifications” like the one above characterize all of Thaeer’s algebraizations. He goes on to say:

About the following four theorems (Dt 17-20) he maintains that they prove that all linear transformations, $\{ax+b|a, b \in \mathfrak{R}\}$ form a group (*dass die ganzen linearen Substitutionen einer Veraenderlichen eine Gruppe bilden*). I am not sure I understand what he means to say, and the *Data* certainly does not help me,—so probably Euclid would not understand either. (p. 77)

Let us take our next example from Archimedes. Against Freudenthal’s assertion, Archimedes’ works are not “instances of algebraic procedure in Greek mathematics.” Heath’s edition of *The Works of Archimedes* (1897) is “in modern notation.” It is faithful only to the disembodied mathematical content of the Archimedean text, but not to its form. And this is crucial. If one abandons Archimedes’ form and transcribes his rhetorical statements by means of algebraic symbols, manipulating and transforming the latter, then

clearly “the algebraic procedure” appears. But this procedure itself is not “in Greek mathematics.” It is a result, as Freudenthal himself states it, of “replacing vernacular by artificial language, and numbering variables by cardinals, a quite recent mathematical tool.” Indeed! Archimedes’ text is anchored securely in the *terra firma* of Greek geometry. If one is not willing to compress wording, to replace “vernacular” by artificial language, to introduce variables and number them by cardinals, and to apply all the other technical tricks which are “quite recent mathematical tools,” then Archimedes’ proof of Proposition 10 of *Peri Helikon* is geometric, not algebraic. This was discerned in a curious way even by Heath, who justified his algebraic procedure and the use of the symbols, “in order to exhibit the *geometrical character of the proof*” (p. 109, my italics).

Dijksterhuis himself in his *Archimedes* said: “In a representation of Greek proofs in the symbolism of modern algebra it is often precisely the most characteristic qualities of the classical argument which are lost, so that the reader is not sufficiently obliged to enter into the train of thought of the original.” So let us oblige ourselves to enter into the train of thought of the original by having a look at the 10th proposition of *Peri Helikon*.

I shall give you the full enunciation and then set at its side Heath’s algebraic variant:

If any number of lines, exceeding one another by the same magnitude, are set one after the other, the excess being equal to the smallest, and if one takes other lines in the same number, each of which is equal to the magnitude of the greatest of the first lines, [then] the squares on the lines equal to the greatest, augmented by the square on the greatest, and by the rectangle the sides of which are the smallest line and the sum of all lines exceeding one another by the same magnitude are equivalent to thrice the sum of the squares on the

lines which exceed one another by the same amount.

The enunciation is geometrical and it is accompanied by simple, linear diagrams.

And here is Heath's enunciation:

If $A^1, A^2, A^3, \dots, A_n$ be n lines forming an ascending arithmetical progression in which the common difference is equal to A_1 , the least term, then

$$(n+1)A_n^2 + A_1(A_1 + A_2 + \dots + A_n) = 3(A_1^2 + A_2^2 + \dots + A_n^2) \dots\dots$$

...the result is equivalent to

$$1^2 + 2^2 + 3^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

The proof is not difficult, but very long (more than three pages in Mugler's edition of Archimedes' works, if one includes the porism), and I think we can dispense with it, but not before pointing out the fact that it lies squarely within the realm of traditional Greek geometry, relying on *Elements* 2.4, it is true, which, as is well known, belongs to the so-called geometric algebra, but which, as I have argued elsewhere, is strictly geometric; additionally, the porism relies on *Elements* 6.20.

As in the enunciation, Archimedes formulates consistently his statements in terms of lines, squares, and rectangles, which he manipulates *à la Grecque*, and his diagrammatic notation is not at all perspicuous to a modern eye, making his transformations opaque, or at least cloudy, to a mind spoiled by the easy mechanics of algebraic manipulations and their immediate visual transparency. This makes following his elementary inferences quite difficult and almost forces upon the reader recourse to algebraic notation. Such a procedure, how-

ever, easy, pellucid, and revealing as it is, is not legitimate historically.

What I have said about the proposition applies in its entirety to the porism following it, out of which Heath makes *two* corollaries!

Let us, again, limit ourselves to the enunciations of Archimedes and Heath, which substantiate our characterization. Archimedes first:

It is manifest from the preceding that the sum of the squares on the lines equal to the greatest is inferior to the triple of the sum of the squares on the lines exceeding one another by the same magnitude, because, if one adds to it [the former] some [magnitudes], it becomes that triple, but that it is superior to the triple of the second sum diminished by the square on the greatest line, since what is added [to the first sum] is inferior to the triple square of the greatest line. It is for this [very] reason that, when one describes similar figures on all the lines, both on those exceeding one another by the same magnitude, as well as on those which are equal to the greatest line, the sum of the figures described on the lines which are equal to the greatest line is inferior to the triple of the sum of the figures described on the lines exceeding one another by the same magnitude, but it is superior to the triple of the second sum, which is diminished by the figure constructed on the greatest line, because similar figures are in the same ratio as the squares [on their sides].

Now Heath:

Cor. 1. It follows from this proposition that

$$nA_n^2 < 3(A_1^2 + A_2^2 + \dots + A_n^2), \text{ and also that}$$

$$nA_n^2 < 3(A_1^2 + A_2^2 + \dots + A_{n-1}^2).$$

Cor. 2. All the results will equally hold if similar figures are substituted for squares.

The differences between Archimedes and Heath are blatant. Faithfulness to the Archimedean way of doing things demands, therefore, the rejection of an edition of his works “edited in modern notation.” There is no escape for the historian but to take texts at their face value.

One last example I shall draw from Diophantus's *Arithmetica*. Diophantus is an exception in the long tradition of Greek mathematics, which is largely geometric. Living probably in the third century A.D., his preserved work is an instance, the only one of its kind in Greek mathematics, of what we call algebra. It is a sui generis algebra, however, in which the notations are properly nonexistent, except for the unknown, *arithmos*, which is itself most likely an abbreviation, and a few abbreviations for powers of the unknown and for the unit, *monas*. That is all. This is what makes his rhetorical algebra syncopated, according to Nesselmann's classification. It is an algebra in which there is a total lack of true, operative symbols, including symbols for operations, relations, and the exponential notation, with the exception of a symbol for subtraction, in which the exceptional skill of Diophantus enables him somehow to overcome with great dexterity the built-in drawbacks of his *Arithmetica*. This Greek algebra, however, is most emphatically not geometric algebra. If anything, it is conceptually a far away relation to what has been traditionally called Babylonian algebra, perhaps not even this, in light of the recent researches of Jens

Hoyrup, who identifies the geometrical roots of the Babylonian recipes [*Lengths, Widths, Surfaces*, (Springer, 2002)].

Whatever it is, it is largely a collection of problems, to be solved by means of skilful guesses and less by systematic methods (though one also finds methods, for example, a general method for the solution of what we call determinate equations of the second degree and another for double equations of the second degree), involving specific known numbers. It is not a book of propositions to be proved, but rather of exercises to be solved, leading mostly to simple determinate and indeterminate equations, by means of which one finds the required numbers, which are always rational, quite often non-integral. Still, despite what I just said, one finds in the treatise also some “porisms” and other *propositions* in the theory of numbers, which proved themselves influential in the history of the theory of numbers, especially in Fermat's work. Thus, Diophantus knew that no number of the form $8n+7$ can be the sum of three squares, and that for an odd number, $2n+1$, to be the sum of two squares, n itself must not be odd, which means that no number of the form $4n+3$ or $4n-1$ can be the sum of two squares.

Now, some simple illustrations from the *Arithmetica*, to get the flavor of the true *Algebra der Griechen*:

Problem 1. To divide a given number into two numbers, the difference of which is known.

Let the given number be 100, and let the difference be 40 monads; to find the numbers.

Let us assume that the smaller number is 1 *arithmos*; hence, the greater number is 1 *arithmos* and 40 units. Therefore, the sum of the two numbers becomes 2 *arithmoi* and 40 units. But the given sum is 100 units; hence, 100 units are equal to 2 *arithmoi* and 40 units. Let us subtract the like from the like, that is, 40 units from 100 and, also,

the same 40 units from the 2 *arithmoi* and 40 units. The two remaining *arithmoi* equal 60 units and each *arithmos* becomes 30 units.

Let us return to what we assumed: the smaller number will be 30 units, while the greater will be 70 units, and the validation is obvious.

And here is Heath's faithful version of the solution:

Given number is 100, given difference 40.

Lesser number required x . Therefore

$$2x+40=100$$

$$x=30.$$

The required numbers are 70, 30.

Problem 2. Diophantus:

It is necessary to divide a given number into two numbers having a given ratio.

Let us require to divide 60 into two numbers in triplicate ratio.

Let us assume that the smaller number is 1 *arithmos*; hence, the greater number will be 3 *arithmoi*, and thus the greater number is thrice the smaller number. It is also necessary that the sum of the two numbers be 60 units. But the sum of the two numbers is 4 *arithmoi*; hence 4 *arithmoi* are equal to 60 units, and the *arithmos* is therefore 15 units. Hence, the smaller number will be 15 units, and the greater 45 units.

Heath:

Given number 60, given ratio 3:1.

Two numbers x , $3x$. Therefore $x=15$.

The numbers are 45, 15.

Finally, an example of what is called indeterminate analysis of the third degree:

Problem 4.8. Diophantus:

To add the same number to a cube and its side and make the same.

Let the number to be added be 1 *arithmos* and the side of the cube be a certain amount of *arithmoi*. Let this amount be 2 *arithmoi* and it follows that the cube is 8 cubic *arithmoi*.

Now if one adds 1 *arithmos* to 2 *arithmoi*, they become 3 *arithmoi*, while if one adds it to the 8 cubic *arithmoi*, they become 8 cubic *arithmoi* and 1 *arithmos*, which are equal to 27 cubic *arithmoi*. Let us subtract 8 cubic *arithmoi*, and it follows that the remaining 19 cubic *arithmoi* will become equal to 1 *arithmos*. Let us divide all by the *arithmos*, and 19 squared *arithmoi* will be equal to 1 unit.

But 1 unit is a square, and if 19, the amount of square *arithmoi*, were a square, the problem would be solved. But the 19 squares find their origin in the excess by which 27 cubic *arithmoi* exceed 8 cubic *arithmoi*; and 27 cubic *arithmoi* are the cube of 3 *arithmoi*, while 8 cubic *arithmoi* are the cube of 2 *arithmoi*. But the two *arithmoi* are taken by hypothesis, and 3 *arithmoi* exceed by one the amount taken arbitrarily as the side. Therefore, we are led to finding two numbers which exceed one another by one unit, and the cubes of which exceed one another by a square.

Let one of those numbers be 1 *arithmos*, and the other 1 *arithmos* plus 1 unit. Hence the excess of their cubes is 3 square *arithmoi* plus 3 *arithmoi* plus 1 unit. Let us set this excess equal to the square the side of which is 1 unit less 2 *arithmoi*,

and the *arithmos* becomes 7 units. Let us return to what we supposed, and one of the numbers will be 7 and the other 8.

Let us now return to the original question and assume the number to be added to be 1 *arithmos*, and the side of the cube to be 7 *arithmoi*. This cube will be then 343 cubic *arithmoi*. Hence, if one adds the *arithmos* to each of these last numbers, one will become 8 *arithmoi* and the other 343 cubic *arithmoi* plus 1 *arithmos*. But we wanted this last expression to be a cube the side of which is 8 *arithmoi*; hence 512 cubic *arithmoi* are equal to 343 cubic *arithmoi* plus 1 *arithmos*, and the *arithmos* becomes $\frac{1}{13}$.

Returning to the things we assumed, the cube will be $\frac{343}{2197}$, the side $\frac{7}{13}$, and the number to be added $\frac{1}{13}$.

I assume you could follow, at least in outline, Diophantus's solution procedure, though this is not essential for my main purpose. (For those of you who could not follow after all the details, which, as I said, is not really necessary, a glance at Heath's, or Ver Eecke's, *Diophantus* would clarify matters.) My purpose is to compare Diophantus to his modern editors. This time, instead of Heath, I shall take, however, Nesselmann.

Nesselmann:

It is necessary that $x^3 + y = (x + y)^3$, i.e., $3x^2 + 3xy + y^2 = 1$.

Solving for y, we get $y = \frac{1}{2} [-3x \pm \sqrt{4 - 3x^2}]$.

Putting $4 - 3x^2 = (2 - \frac{m}{n}x)^2$, one gets $x = \frac{4mn}{3n^2 + m^2}$.

Hence, $y = \frac{-6mn \pm (m^2 - 3n^2)}{3n^2 + m^2}$. Taking only the + sign,

for y to be positive, it is necessary that $m^2 - 3n^2 > 6mn$,

or, $\frac{m^2}{n^2} - 6\frac{m}{n} > 3$, or, $(\frac{m}{n} - 3)^2 > 12$, or $(\frac{m}{n} > 3 + \sqrt{12}$.

Diophantus's solution corresponds to $m = 7$, $n = 1$.

Now this is historical faithfulness!

Time to conclude. Historians must take the past seriously. For historians of science and mathematics, this means taking texts seriously. How does one do this? By reading them as they are, in their nakedness, as it were, in the language in which they were written, without shortcuts and transmogrification, resulting in their translation into scientific and mathematical languages which became historically available only long after they were written. It is, therefore, crucial that the form of those texts remain inviolable. Without this, anachronism, i.e., historical misunderstanding, becomes rampant and the resulting interpretation is misinterpretation. As Benjamin Farrington put it,

History is the most fundamental science, for there is no human knowledge which cannot lose its scientific character when men forget the conditions under which it originated, the questions which it answered, and the function it was created to serve. A great part of the mysticism and superstition of educated men consists of knowledge which has broken loose from its historical moorings.

In 1975 an article appeared in the *Archive for History of Exact Sciences*, arguing for the need to rewrite the history of Greek mathematics. In one of its many footnotes, namely the one numbered 126, attention is called to an important book, *Greek Mathematical Thought and the Origin of Algebra*. The footnote in question contains a parenthesis, saying:

[L]et me urge those readers who have a choice and wish to read [this] highly interesting study to refer back to the original German articles: somehow the

pomposity, stuffiness, and turgidity of the author's style are better accommodated by the Teutonic cadences than by the more friendly sounds of the perfidious Albion.

This was a nasty and uncalled for remark, attributable to the author's hubris.

How pleasantly and embarrassingly surprised must he have been, then, when, a few weeks after the article's appearance in the *Archive*, a postcard from the slighted author of the book (to whom no reprint was sent!) arrived, saying: "Dear Dr.,... Thank you very much for your important article in the *Archive*...." In character, you may say. Indeed.

As a belated, and highly inappropriate atonement for that unsavory footnote, its author would like to finish this lecture with a highly pertinent and lengthy quotation from an article written by the great scholar he so carelessly slighted:

[T]he relation between ancient and modern mathematics has increasingly become the focus of historical investigation... Two general lines of interpretation can be distinguished here. One—the prevailing view—sees in the history of science a continuous forward progress interrupted, at most, by periods of stagnation. On this view, forward progress takes place with 'logical necessity,' accordingly, writing the history of a mathematical theorem or of a physical principle basically means analyzing its logic. The usual presentations, especially of the history of mathematics, picture a rectilinear course; all of its accidents and irregularities disappear behind the logical straightness of the whole path.

The second interpretation emphasizes that the different stages along this path are incomparable....it sees in Greek mathematics a science totally distinct from modern mathematics....*Both* interpretations, however, start from the present-day

condition of science. The first measures ancient by the standard of modern science and pursues the individual threads leading back from the valid theorems of contemporary science to the anticipatory steps taken towards them in antiquity.... The second interpretation strives to bring into relief, not what is common, but what divides ancient and modern science. It too, however, interprets the otherness of ancient mathematics...in terms of the results of contemporary science. Consequently, it recognizes only a counter-image of itself in ancient science, a counter-image which still stands on its own conceptual level.

Both interpretations fail to do justice to the true state of the case. There can be no doubt that the science of the seventeenth century represents a direct continuation of ancient science. On the other hand, neither can we deny their differences...*above all*, in their basic initiatives, in their whole disposition (*habitus*). The difficulty is precisely to avoid interpreting their differences and their affinity one-sidedly in terms of the new science. The issues at stake cannot be divorced from the specific conceptual framework within which they are interpreted....

We need to approach ancient science on a basis appropriate to it, a basis provided by that science itself. Only on this basis can we measure the transformation ancient science underwent in the seventeenth century—a transformation unique and unparalleled in the history of man....

This modern consciousness is to be understood not simply as a linear continuation of ancient *επιστημη*, but as the result of a fundamental conceptual shift which took place in the modern era, a shift we can nowadays scarcely grasp.

The name of the man who wrote these lines is, as I am sure you know, Jacob Klein. *Yehi Zichro Baruch!* May his memory be blessed!



A Note on the Opposite Sections and Conjugate Sections in Apollonius of Perga's *Conica*

Michael N. Fried

Introduction

To a careful reader of Apollonius of Perga's *Conica*, the difference between Apollonius' view of conic sections and ours ought to be evident on nearly every page of the work. Yet this has not always been the case. Indeed, it has not always been easy to persuade readers that there really is something Greek about Apollonius' mathematics. Partly to blame is the seductive power of the algebraic framework in which we study conic sections today and in which historians of mathematics have interpreted the book in past years.¹ Viewed algebraically, for example, all of Book 4 of the *Conica*—the book concerning the number of points at which conic sections can meet—can be reduced to a single proposition, namely, that a system of two quadratic equations in two unknowns can have at most four solutions. This tremendous power allowing one to obtain results corresponding to ones in the *Conica* makes it all too easy to think that Apollonius' own text, in effect, can be bypassed and replaced by an updated algebraic version of it. However, there are things in the *Conica* that are refractory to this kind of modernization of the text and show its essentially geometric character. One of these, surely, is that most peculiar entity in the *Conica*, the opposite sections; it is they that I shall turn my attention to in this note. In particular, I want to show something about status of the opposite sections in the *Conica* and show, among other things, why their status,

Michael Fried teaches mathematics education at Ben-Gurion University of the Negev. He is the author of a translation, with introduction, of *Book Four of the Conics of Apollonius*, and is the co-author (with Sabetai Unguru) of *Apollonius of Perga's Conica: Text, Context, and Subtext*.

why their place in the *Conica*, is different from that of the conjugate sections.

The Absence of the Opposite Sections from the Modern View

The reason why looking at the opposite sections is a good way to re-focus on Apollonius' text, rather than on an algebraic reconstruction of it, is simply that the opposite sections do not exist in modern mathematics. For us, there is only the hyperbola. Our view of the hyperbola, on the other hand, is that it is a curve consisting "of two open branches extending to infinity."² That we are not bothered by one curve consisting of two is directly related to our defining the hyperbola, as we do the other conics, by means of an equation. Defining it this way, the hyperbola becomes merely a set of points given by coordinates, say the Cartesian coordinates (x, y), satisfying an equation such as this:

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

Hence, there is nothing shocking when we discover that this set of points consists of two disjoint subsets, one containing points whose x coordinates are less than or equal to $-a$ (where a is taken to be a positive real number) and one containing points whose x coordinates are greater than or equal to $+a$; the only relevant question to ask is whether the coordinates of a given point satisfy or do not satisfy the given algebraic relation.³ The equation, in this view, tells all; it contains all the essential information about the object; in a sense, the equation of the hyperbola *is* the hyperbola (see Fried & Unguru, 2001, pp. 102-103; Klein, 1981, pp. 28-29). To speak about opposite sections in addition to the hyperbola *is* unnecessary because they are not distinguished by different equations.

The Opposite Sections in the *Conica*

The point of view above has also been the point of departure for an older, but still much listened to, generation of historians of mathematics. Zeuthen (1886) and Heath (1921, 1896), leaders of that generation, had no doubt that Apollonius' achievement was in the uncovering and elucidation of the equations of the conic sections and that Apollonius understood the conic sections in an algebraic spirit. For them, Apollonius' view was the modern view. Thus Zeuthen writes:

An ellipse, parabola or hyperbola is here [in Book 1] planimetrically determined as a curve which is represented by the equation (3), (1) or (2) [the Cartesian equations for the ellipse, parabola, and hyperbola, respectively] in a system of parallel coordinates with any angle between the axes. Thus, apart from the determination of the position of these curves, they seem to depend on three constants, namely, that angle [between the axes, or, equivalently, the ordinate angle], p [*latus rectum*], and a [the length of the diameter] (Zeuthen, 1886, p. 67).

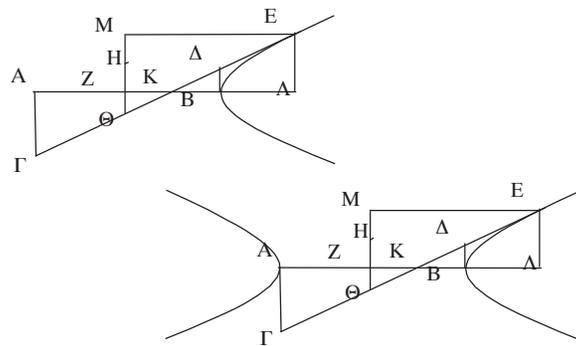
And Heath, echoing Zeuthen, writes:

Apollonius, in deriving the three conics from any cone cut in the most general manner, seeks to find the relation between the coordinates of any point on the curve referred to the original diameter and the tangent at its extremity as axes (in general oblique), and proceeds to deduce from the relation, when found, the other properties of the curves. His method does not essentially differ from that of modern analytical geometry except that in Apollonius geometrical operations take the place of algebraic calculations (Heath, 1896, p. cxvi)

As for why Apollonius notwithstanding referred to both opposite sections *and* hyperbolas, Heath remarks that, “*Since* [Apollonius] *was the first to treat the double-branch hyperbola fully* [emphasis added], he generally discusses the *hyperbola* (i.e., the single branch) [Heath’s emphasis] along with the ellipse, and *the opposites* [Heath’s emphasis], as he calls the double-branch hyperbola, separately” (Heath, 1921, p.139). Zeuthen simply says that, as a matter of terminology, “What [Apollonius] calls an hyperbola is always only a hyperbola-branch” (Zeuthen, 1886, p. 67). But these remarks are hardly satisfying. First, while it is most likely that Apollonius was truly the first to treat the “double-branch hyperbola”—that is, the opposite sections—fully, the opposite sections were not so completely unfamiliar to his contemporaries that he had to continually remind them of their existence; indeed, in the preface to Book 4 he refers to the opposite sections as if they were known and discussed, at least by Conon of Samos and Nicoteles of Cyrene. Second, although Zeuthen’s remark is not meant to explain Apollonius’ use of both terms, it still begs the question: if what Apollonius called the hyperbola is truly represented by the Cartesian equation and therefore, is the modern double-branched hyperbola (as Zeuthen unambiguously implies throughout his book), why cause confusion by referring to a branch of the hyperbola as the hyperbola and the two branches together as something else, namely, the opposite sections? Zeuthen himself never speaks of “opposite sections” but only of the “two (or “corresponding”) branches of the hyperbola” the “whole hyperbola” (*vollständige Hyperbel*), and, most often, simply the “hyperbola.” In this way, he and his followers, including Heath, merely push aside as if nugatory the obvious fact that in the *Conica*, from start to finish, there are hyperbolas and there are opposite sections.

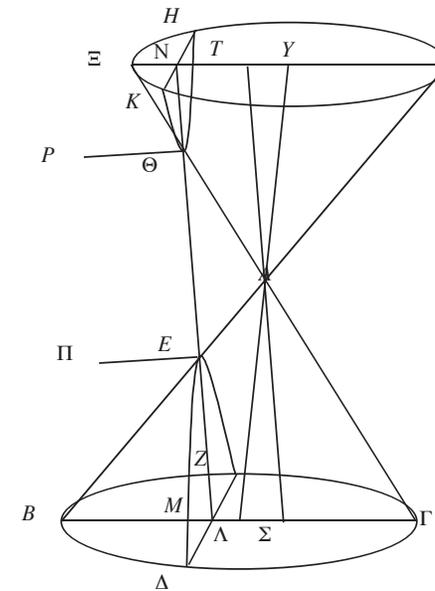
The opposite sections (*hai tomai antikeimenai*) are quite prominent in the work. They figure in 24 of the 53 propositions in Book 2; 41 of the 56 propositions in Book 3, and 38 of the 57 propositions in Book 4. In all these propositions, Apollonius treats opposite sections as something apart from

hyperbolas. A sign of this is that the opposite sections are separated from the hyperbola in the enunciations of propositions. For example, the statement of proposition 3.42 is: “If in an hyperbola or ellipse or circumference of a circle or opposite sections [emphasis added] straight lines are drawn from the vertices of the diameter parallel to an ordinate, and some other straight line at random is drawn tangent, it will cut off from them straight lines containing a rectangle equal to the fourth part of the figure to the same diameter.” 3.44 reads, “If two straight lines touching an hyperbola or opposite sections [emphasis added] meet the asymptotes, then the straight lines drawn to the sections will be parallel to the straight line joining the points of contact.” If one knew nothing about the opposite sections, these statements would suggest that the hyperbola and the opposite sections were as different from one another as the hyperbola is different from the ellipse.⁵ The proofs of these two particular propositions make little distinction between the hyperbola and opposite sections. In 3.44 a separate diagram for the opposite sections is needed to bring out a case that applies to the two sections together, but in 3.42 not even that is needed, despite Heiberg’s insistence on adding an extra diagram for the opposite sections anyway (fig. 1). Nothing from the logical rigor would be lost if Apollonius referred only to the opposite sections in these propositions. So even where the logic does not require it, Apollonius makes a point to separate, in words verbally, the hyperbola from the opposite sections.

Fig. 1: *Conica* 3.42 (Heiberg)

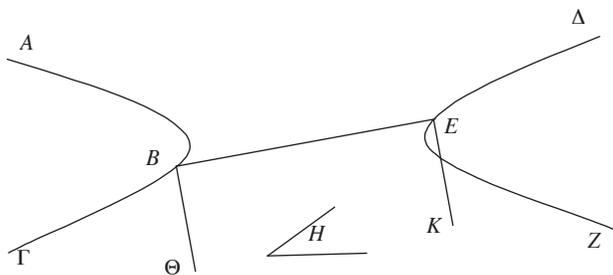
Of course this is not to say there is no connection between the opposite sections and the hyperbola, for the central property of the opposite sections is that they are composed of two hyperbolas. Apollonius proves this in the proposition that introduces the opposite sections, proposition 1.14:

If the vertically opposite surfaces are cut by a plane not through the vertex [see fig. 2], *the section on each of the two surfaces will be that which is called the hyperbola* [emphasis added]; and the diameter of the two sections will be the same straight line; and the straight lines, to which the straight lines drawn to the diameter parallel to the straight line in the cone's base are applied in square, are equal; and the transverse side of the figure, that between the vertices of the sections, is common. And let such sections be called opposite (*kaleisthōsan de hai toiautai tomai antikeimenai*)

Fig. 2: Diagram for *Conica* 1.14

This is also the way he refers to the opposite sections later, especially in Book 4, where he consistently speaks of “an hyperbola and its opposite section.”⁷ The genitive in those phrases suggests that the opposite section belongs to the given hyperbola, that is, every hyperbola has its very own opposite section.

The most striking instance in which Apollonius refers to the opposite sections as two hyperbolas is in his construction of the opposite sections in 1.59. It is worth reviewing how this construction is carried out. Apollonius asks, specifically, for the following:

Fig. 3: Diagram for *Conica*, 1.59

Given two straight lines perpendicular to one another, find opposite [sections], whose diameter is one of the given straight lines and whose vertices are the ends of the straight line, and the lines dropped in a given angle in each of the sections will [equal] in square [the rectangles] applied to the other [straight line] and exceeding by a [rectangle] similar to that contained by the given straight lines.⁸

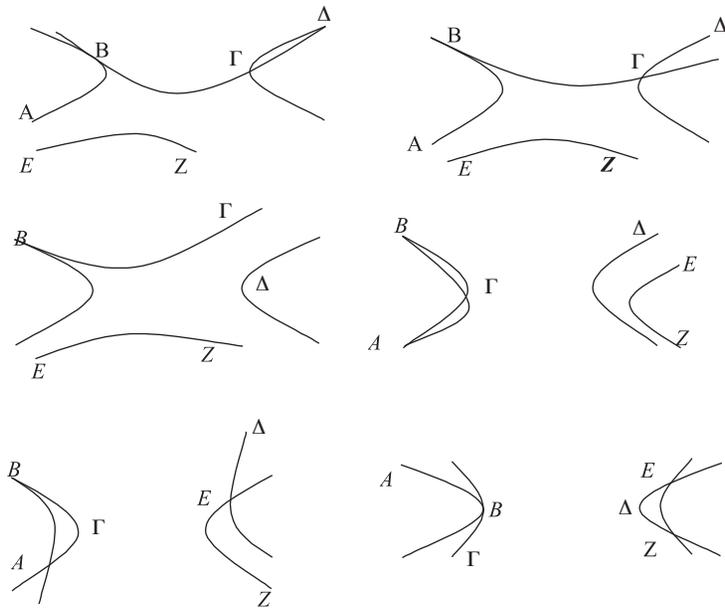
Let BE and $B\Theta$ be the given lines and let H be the given angle (see fig. 3). Choosing BE to be the transverse diameter (*plagia*) and $B\Theta$ to be the upright side of the figure (*orthia*), Apollonius says to construct an hyperbola $AB\Gamma$, adding, that “This is to be done as has been set out before (*prosgegraptai*).” For the latter, he has in mind 1.54-55 where he shows how an hyperbola is constructed having a given diameter, upright side, and ordinate angle. Next he says: let EK have been drawn through E perpendicular to BE and equal to $B\Theta$, and draw an hyperbola ΔEZ having BE as its diameter, EK its upright side, and H its ordinate angle, again, presumably, relying on 1.54-55. With that, he concludes: “It is evident (*phaneron dē*) that B and E [i.e., $AB\Gamma$ and ΔEZ] are opposite [sections] and they have one diameter and equal upright sides.” I shall return to the question whether it truly is evident that B and E are opposite sections, but for now what is important to understand is that Apollonius constructs the

opposite sections by constructing one hyperbola and then another.

Hence the opposite sections are two hyperbolas but not any two hyperbolas; they are two hyperbolas somehow linked together; they belong to one another the way identical twins do.⁹ In a sense, then, Apollonius, as I understand him and as Zeuthen understands him, begins with the hyperbola; however, we disagree about the direction in which he proceeds. Whereas Zeuthen begins with the hyperbola as the two-branched curve given by the Cartesian equation and then focuses on the single branch, which Apollonius calls the hyperbola, my view is that Apollonius begins with the single connected curve, which he calls the hyperbola, and then investigates the opposite sections in terms of it. Although Ockham’s razor would probably fall on the side of the latter view, we need to get a better grasp of this peculiar situation wherein two well-defined curves, the two hyperbolas, become together a distinct geometrical entity, the opposite sections. To do this, let us consider the ways in which two conic sections are juxtaposed in the *Conica*, that is to say, let us consider how Apollonius treats pluralities of curves.

Pluralities of Curves

Surely, the juxtapositions of conic sections are to be found in Book 4 of the *Conica*. The very point of that book is to investigate the ways in which conic sections can come together—whether they touch, whether they intersect, at how many points can they touch, at how many they can intersect, and so on. The variety of configurations Apollonius considers can be seen in the diagram for, say, 4.56:

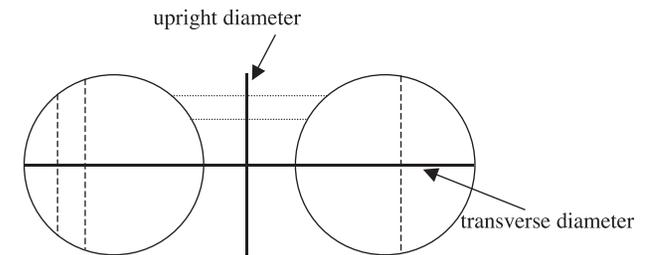
Fig.4: Diagram for *Conica* 4.56

As discussed elsewhere,¹⁰ one of the fundamental facts explored in Book 4 is the ability of conic sections to be placed arbitrarily in the plane, as Euclid postulates for circles and lines—a fact partially justifying Book 4’s inclusion into what Apollonius terms a “course in the elements of conics.” But it is this arbitrariness that makes the question of the plurality of conic sections in Book 4 the exact opposite of the one we are trying answer about the opposite sections; for the opposite sections are not thrown together; they belong together.

The first indication that a pair of curves may be associated with one another as the opposite sections is given by Apollonius in the definitions at the start of Book 1. Among these definitions are those for the *transverse* and *upright* diameters (*diametros plagia* and *diametros orthia*). Having defined the diameter of a curved line, Apollonius continues, “Likewise, of any two curved lines (*duo kampilôn grammôn*) lying in one plane, I call that straight line the transverse diameter which cuts the two curved lines and bisects all the

straight lines drawn to either of the curved lines parallel to some straight line...and I call that straight line the upright diameter which, lying between the two curved lines, bisects all the straight lines intercepted between the curved lines and drawn parallel to some straight line.” An illustration for the upright and transverse diameter can be seen in fig. 5 below, which is not a diagram from Apollonius’ text.

Fig. 5 The upright and transverse diameters

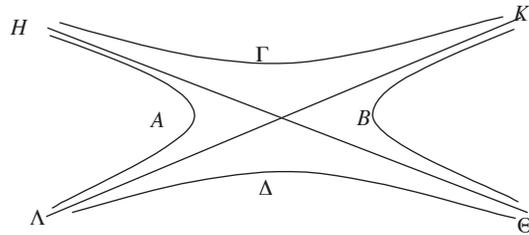


These definitions prepare us for the transverse diameter of the opposite sections introduced in the same proposition in which the opposite sections themselves are defined, 1.14; that the two hyperbolas making up the opposite sections have a common diameter is one of the things that links them together. Interestingly enough, the upright diameter, which is the diameter explicitly defined for two curves and that which is most clearly applicable to the opposite sections,¹¹ is rarely used by Apollonius: it is used twice in Book 2 (2.37, 38) and once in Book 7 (7.6); and even in those cases Apollonius is quick to identify it as one of a pair of *conjugate diameters* (*suzugeis diametroi*), which are good for *both* two curves and one. So while Apollonius provides for the possibility of distinct curves linked together, he seems almost to avoid the idea, at least with regards to the opposite sections.

The Conjugate Sections

But the opposite sections are not the *only* curves linked together in the *Conica*. What Apollonius calls the conjugate sections (*hai tomai suzugeis*) consist of four hyperbolas, or, to be precise, two pairs of opposite sections.¹² The very word

Fig.6: Conjugate Sections



suzugeis refers to a couple yoked together, a married pair.¹³ What links them together? First, by the definition given them in 1.60, the diameter of the one pair is the conjugate diameter of the other, that is, if the diameters of AB and ΓΔ are D and d , respectively, then lines drawn in AB parallel to d will be bisected by D , and *vice versa*; moreover, D is equal in square to the rectangle contained by d and the *latus rectum* with respect to d (this rectangle being the “figure” or *eidos* of the opposite sections), while d is equal in square to the rectangle contained by D and the *latus rectum* with respect to D . Second, as Apollonius shows in 2.17, “The asymptotes [lines $H\Theta$ and AK in fig. 6] of conjugate opposite sections are common.” This is also shown for the opposite sections in 2.15; these shared asymptotes are certainly a crucial unifier not only of the conjugate sections but also of the two opposite sections themselves.¹⁴

Compared to the arbitrary juxtapositions of conic sections presented in Book 4, the opposite sections and the conjugate sections seem to be quite similar in that their component curves are bound by means of asymptotes and diameters. Indeed, Apollonius often treats the opposite sections and conjugate sections in analogous propositions for instance: 2.41 states, “If in opposite sections two straight lines not through

the center cut each other, then they do not bisect each other”; and 2.42, states that “If in conjugate opposite sections two straight lines not through the center cut each other, they do not bisect each other.” Yet this close connection between the opposite sections and conjugate sections raises the question about why the opposite sections nevertheless enjoy a status in the *Conica* that the conjugate sections do not? Why, in particular, are the conjugate sections never mentioned in conjunction with the other conic sections as are the opposite sections? Why are only the opposite sections included in the clique, parabola, hyperbola, ellipse, and opposite sections?¹⁵ For this, we must return to the beginning of the *Conica*, for the answer has very much to do with beginnings.

The Genesis of the Opposite Sections

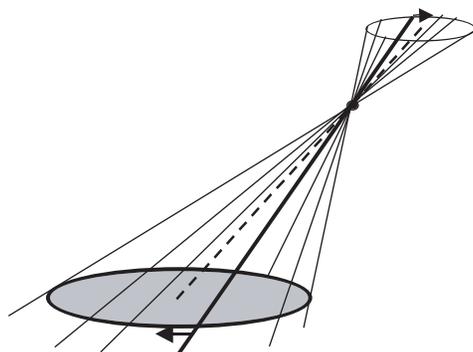
The *Conica* opens not with the definition of the conic sections, but with the definition of the conic surface (*kônikê epiphaneia*) from which arises the figure of the cone. Here is Apollonius’ definition:

If from a point a straight line is joined to the circumference of a circle which is not in the same plane with the point, and the line is produced in both directions, and if, with the point remaining fixed, the straight line being rotated about the circumference of the circle returns to the same place from which it began, then the generated surface composed of the two surfaces lying vertically opposite one another, each of which increases indefinitely as the generating straight line is produced indefinitely, I call a conic surface, and I call the fixed point the vertex.....And the figure contained by the circle and by the conic surface between the vertex and the circumference of the circle I call a cone.

The definition is vivid and visual owing in large part to its motion-imbued language. Motion, as is well known, is

avoided in Greek mathematics. Thus, much has been made of the hints of motion in *Elements*, 1.4. But avoiding motion seems to be a desideratum chiefly for propositions and demonstrations, for motion in *definitions* is not all that unusual. Besides this one from the *Conica*, Archimedes gives similar kinematic definitions for conoids and spheroids in the letter opening *On Conoids and Spheroids*, and Euclid himself defines the cone, sphere, and cylinder in Book 11 of the *Elements* by means of rotating a triangle, semicircle, and rectangle, respectively. The presence of motion in these definitions gives one the sense that one is witnessing the very coming to be of the objects being defined. In this way, such definitions take on a mythic quality; one almost wants to see the *demiourgos* turning the generating line about the circumference of the base circle.

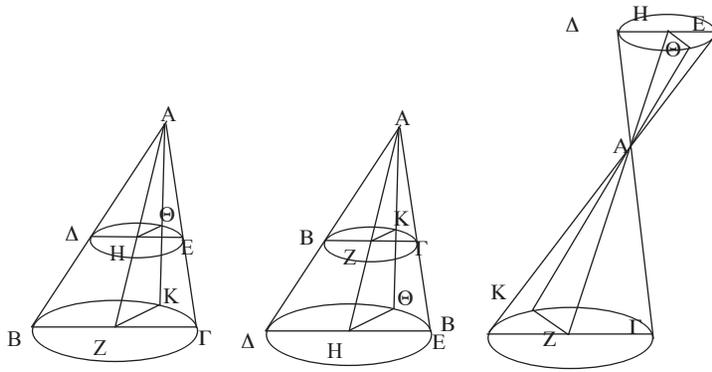
Fig.7: The Conic Surface



Following the definitions, Apollonius proceeds in ten propositions to develop the basic properties of sections of cones and, in particular, to show (in 1.7) how a cone may be cut so that the section produced will be endowed with a diameter. When the reader arrives at 1.11, which begins the series of four propositions defining the parabola, hyperbola, ellipse, and opposite sections, one is ready to see how these sections arise within the cone. I do not use the word “see” lightly. In these propositions, Apollonius dedicates considerable space both in the enunciation and in the body of the

proof to describe the sequence of geometrical operations undertaken to produce the particular sections; just as one witnesses the coming to be of the conical surface and cone, now one witnesses the coming to be of the conic sections.

Of all the sections presented in 1.11-14, the one that is most obviously related to Apollonius’ cone is in fact the opposite sections. The reason is clear when one considers Apollonius’ definition of the conical surface and cone. When Euclid, by contrast, defines a cone in Book 11 of the *Elements*, he does it by describing the rotation of a right triangle about one of its legs. Hence, Euclid begins by defining the figure of the cone; moreover, Euclid’s cone, from the start, is right, that is, its axis is perpendicular to its base, and it is bounded. Apollonius begins by defining a conic surface, which is generally oblique (since the line from the fixed point to the center of the circle about whose circumference the generating line turns is not necessarily perpendicular to the plane of the circle), unbounded, and, significantly, double. The cone arises from this surface as the figure (*schema*) contained by the base circle and the vertex. Later, in proposition 1.4, Apollonius shows that by cutting the conic surface with planes parallel to that of the base circle any number of cones may be produced from the conical surface. In this way, the two vertically opposite surfaces (*hai kata koruphên epiphaneiae*) of the conic surfaces give rise to cones on either side of the vertex.

Fig. 8: *Conica* 1.4

The doubleness of the conic surface and the double set of cones that arise from it is (together with its obliqueness) surely one of the most striking aspects of Apollonius' definition. So, when Apollonius defines the opposite sections in 1.14, the definition is not entirely unexpected; it has been prefigured in the definition of the conic surface itself. Indeed, while the hyperbola is defined in terms of the figure of the cone, which is defined by means of the conic surface, the opposite sections are defined in terms of the conic surface, and in this sense, the opposite sections are prior to the hyperbola. It is worth observing in this connection that 1.14 is the last proposition in the book in which the conic surface appears, as if to say that it no longer has to appear, having served its purpose.

But here a little more needs to be said. For while it is true that the conic surface does not appear again in the *Conica*, the cone does: it reappears in Book 6, where it is used to carry out various constructions connected with similar and equal conic sections, and, most importantly, at the end of Book 1, where Apollonius constructs conic sections in a plane having given diameters, *latera recta*, and ordinate angles. In the latter constructions, Apollonius generally proceeds by taking the given plane as the cutting plane and then constructing the cone cut by that plane so that the section pro-

duced is the section required. One might expect, therefore, that when in 1.59 he set out to construct opposite sections, he would similarly have constructed the conic surface cut by the given plane to produce opposite sections. But, as we saw above, this is not what he does. Applying 1.54-55 in the same sequence, he constructs one hyperbola and then another. Furthermore, if one follows 1.54-55 to the letter in constructing the hyperbolas on both sides of BE, which involves constructing cones as I have remarked, one does not arrive at the two vertical opposite surfaces composing the conic surface without significantly altering the construction in 1.54-55. Can Apollonius truly say, then, as he does, that "it is evident" (*phaneron de*) that opposite sections are produced in 1.59? I think he can. For when he produces the first hyperbola in 1.59, he does so by producing a cone, in accordance with 1.54-55; but the cone is only a figure cut off from a conic surface, and we know from 1.14 that when the plane of the first hyperbola cuts the opposite surface of the conic surface it will produce the opposite section of the first hyperbola. What remains for Apollonius is only to produce that second hyperbola so that it has the right diameter, *latus rectum*, and ordinate direction, which is precisely what he does.¹⁶ So, although the conic surface does not appear explicitly in 1.59, it is still there implicitly. In fact, this is the case with all the constructions at the end of Book 1: the initial construction of the parabola, ellipse, and hyperbola, in which the ordinate direction is right, is carried out by explicitly constructing a cone, but in the continuation of the constructions, that is, in the cases in which the ordinate directions are not right, the cone does not appear. In this way, I think it is wrong to think Apollonius is trying to break away from the cone; the conic sections are *rooted* in the cone. Indeed, it is like the roots of a tree: although the roots are unseen, the leaves, even those at the very summit of the tree, will not survive without them.

Thus we can still say with confidence that the opposite sections have the special status that they do because of their immediate origin in the double conic surface. They inherit

their doubleness directly from the doubleness of the conic surface; their unity they obtain from the single plane that cuts the surface. This geometric birth of the opposite sections sets them apart from the conjugate sections and puts them in the same class as the parabola, ellipse, and hyperbola.

Summary and Concluding Remarks

Apollonius' treatment of the opposite sections tells us much about his mathematical world—mostly because these sections are not found outside of it. In modern mathematics there is only the hyperbola, and the hyperbola has two branches. In Apollonius' geometry there is an hyperbola and there are opposite sections. The hyperbola is produced by cutting a cone; it has an opposite section since the conic surface from which the cone arises extends not only below but also above the vertex of the cone. The opposite sections are linked by their common asymptotes and common diameter. The conjugate sections are linked similarly by common asymptotes and, though not a common diameter, by conjugate diameters. However, this kind of linkage seems to be less compelling for Apollonius than the linkage arising from the single plane cutting the conic surface, for only the opposite sections, and not the conjugate sections, are spoken of in conjunction with the other conic sections.

Earlier I referred to Apollonius' definition of the conic surface as having a mythic quality. This was to give some explanation of the motion-filled description constituting the definition and to highlight the possibility that the development of the conic sections from the conic surface and the figure of the cone was, for Apollonius, nothing short of a spectacle of mathematical genesis. The use of the word "myth" in a mathematical context is jarring for modern ears; mathematics, if to anything, should be related to *logos* not *muthos*. But in the Greek world, it must be recalled, *muthos* and *logos* overlapped as well as being opposed (Peters, 1967, pp. 120-121), so that relationship between the two was one of tension rather than exclusion. Cassirer has gone far to show that the

myth-making function is fundamental to human experience, and I rather agree with him when he says:

The mythical form of conception is not something superadded to certain definite *elements* of empirical existence; instead, the primary "experience" itself is steeped in the imagery of myth and saturated with its atmosphere. Man lives with *objects* only in so far as he lives with these forms; he reveals reality to himself, and himself to reality in that he lets himself and the environment enter into this plastic medium, in which the two do not merely make contact, but fuse with each other. (Cassirer, 1946, p. 10)

Be that as it may, the use of the word "myth" in the context of the *Conica* does suggest a view of Apollonius as one very much engaged with the being of his objects and where they come from; indeed, the distance we feel between myth and mathematics is partly the result of an overly pragmatic view of mathematics in which properties take precedence over origins, a view which is characteristically modern. In this paper, I have tried to show that if we disregard visible geometrical origins and focus only on abstract relations, such as are captured in an algebraic equation, it becomes difficult to understand why the opposite sections enjoy the status they do in the *Conica* as conic sections different from the hyperbola, and why their status is different from that of the conjugate sections—and, conversely, I have tried to show that the opposite sections bring us back to the importance of geometrical origins in the *Conica* and, in so doing, allow us one key to the character of classical mathematics.

Bibliography

Apollonius. (1891, 1893). *Apollonii Pergaei quae Graece exstant cum commentariis antiquis*. Edited by I. L. Heiberg. 2 volumes. Leipzig: Teubner.

Apollonius. (1998). *Apollonius of Perga Conics Books I-III*. Translated by R. Catesby Taliaferro. Santa Fe: Green Lion Press.

Apollonius. (2002). *Apollonius of Perga Conics Book IV*. Translated with introduction and notes by Michael N. Fried. Santa Fe: Green Lion Press.

Cassirer, E. (1946). *Language and Myth*. New York: Harper and Brothers.

Fried, Michael N. and Unguru, Sabetai. (2001). *Apollonius of Perga's Conica: Text, Context, Subtext*. Leiden: Brill.

Goodwin, W. W. & Gulick, C. B. (1992). *Greek Grammar*. New Rochelle: Aristide D. Caratzas.

Heath, Thomas L. (1896). *Apollonius of Perga, Treatise on Conic Sections*. Cambridge: At the University Press.

Heath, Sir Thomas L. (1921). *A History of Greek Mathematics*, 2 vols., Oxford: At the Clarendon Press (repr., New York: Dover Publications, Inc., 1981).

Klein, Jacob (1981). "The World of Physics and the Natural World." *The St. John's Review*, 33 (1): 22-34

Peters, F. E. (1967). *Greek Philosophical Terms*. New York: New York University Press.

Sommerville, D. M. Y. (1946). *Analytic Conics*. London: G. Bell and Sons, Ltd.

Toomer, G. J. (1990). *Apollonius Conics Books V to VII: The Arabic Translation of the Lost Greek Original in the Version of the Banu Musa*. 2 vols. (Sources in the History of Mathematics and Physical Sciences, 9), New York: Springer-Verlag

Toomer, G. J. (1970). Apollonius of Perga, in *DSB*, 1, 179-193.

Youschkevitch, A. P. (1976). The Concept of Function up to the Middle of the 19th Century. *Archive for History of Exact Sciences* 16, pp. 37-88.

Zeuthen, H. G. (1886). *Die Lehre von den Kegelschnitten im Altertum*, Kopenhagen, (repr., Hildesheim: Georg Olms Verlagsbuchhandlung, 1966).

Notes

¹ The algebraic reading of the *Conica* as an historical interpretation has, of course, Zeuthen (1886) as its greatest representative; indeed, Zeuthen's view of the *Conica* was completely adopted by Heath (1896) and, to some extent, by Toomer as well (Toomer, 1990, 1970).

² From Sommerville (1946), but, needless to say, a similar description can be found in any other textbook of analytic geometry.

³ In this connection, it is worth recalling that the sense in which functions such as $f(x) = a/x$ (whose graph, of course, is also an hyperbola) are to be considered *continuous* or not has itself an interesting history, one that shows again how history rarely moves in straight lines. For Euler, "continuity" referred to the rule, so that a function, like $f(x) = a/x$, governed by a single analytic equation should called "continuous." Although this view still echoes in our calling the hyperbola *one* curve with two branches, Euler's approach to the continuity of functions gave away to a more geometrical view of continuity, that is, where continuity has much to do with connectedness in the work Cauchy and Bolzano (For a detailed discussion of these shifts in the meaning of continuity in the history of the function concept, see Youschkevitch, 1976).

⁴ They are less prominent, however, in the extant later books, Books 5-7 and in Book 1, where they appear in only 9 propositions.

⁵ The same argument also shows the difficulty in maintaining that, for Apollonius, the circle was a kind of ellipse. The relationship between the ellipse and the circle has been discussed extensively in Michael N. Fried & Sabetai Unguru, *Apollonius of Perga's Conica: Text, Context, Subtext*, chap. 7.

⁶ Unless stated otherwise, all translations from Books 1-3 are from R. Catesby Taliaferro in *Apollonius of Perga Conics: Books I-III* (Green Lion edition).

⁷ For example in 4.48 (although any proposition from among 4.41-54 could be taken as an example) we have in the *ekthesis*: *estôsan antikeymenai hai ABΓ, Δ, kai huperbole tis he AHΓ...kai tês AHΓ antikeymene estô he E*.

⁸ My translation.

⁹ For this reason, one might expect the dual to be used in referring to the opposite sections instead of the simple plural. The dual was used for denoting “natural pairs” (see W. W. Goodwin & C. B. Gulick, *Greek Grammar*, §§ 170, 838, 914). But, as Christian Marinus Taisbak has pointed out to me, it seems that by Hellenistic times “the dual had become more or less obsolete.”

¹⁰ Fried & Unguru, *op. cit.*, chap. 3; Fried, M. N., *Apollonius of Perga: Conics Book IV*, pp. xxi-xxvii.

¹¹ Thus, it is not at all surprising that the figure Eutocius uses to illustrate Apollonius’ general definition of the transverse diameter is evidently a pair of opposite sections (Heiberg, 2.201), and not, say, a pair of circles, which would have served as well and which I have used above.

¹² The conjugate sections are introduced in the last proposition of Book 1, proposition 60. Besides 1.60, the conjugate sections appear also in propositions 2.17-23,42-43; 3.13-15,23-26,28-29, as well as in a very striking proposition in Book 7, 7.31.

¹³ For the verb *suzeugnumi*, from which *suzugeis* is derived, Liddell and Scott give the definitions, “to yoke together, couple or pair...esp. in marriage.” It is worth noting that in 4.49, 50, 51 Apollonius uses the same adjective *suzugeis* to refer to opposite sections.

¹⁴ See Fried & Unguru, *op.cit.*, pp. 123-124.

¹⁵ In Book 4, particularly in 4.25 (which is the central theorem of the book: “A section of a cone does not cut a section of a cone or circumference of a circle at more than four points”) the opposite do seem to be separated from the other conic sections. In this specific case, however, Heath’s argument above might be correct, namely, that the cases involving the opposite sections, as Apollonius stresses in the preface, were new and needed to be highlighted. It may also be that 4.25 does actually intend the opposite sections, even though the proof for the opposite sections comes latter; that kind of nexus is not unusual in the *Conica*. Whatever the case, it still remains that, unlike the opposite sections, the *conjugate sections* do not appear at all in Book 4.

¹⁶ One might argue here that he is using a result here from Book 6, namely, that two hyperbolas having the equal and similar figures are

equal (6.2), but this could be argued regarding all the constructions at the end of Book 1.



Viète on the Solution of Equations and the Construction of Problems

Richard Ferrier

Introduction

I wish to do two things in this paper, first to review the grounds for taking François Viète as a pivotal figure in the history of thought, and second to clarify an interesting technical aspect of his work, namely, what he called the “exegetical art.” This clarification will show Viète to be a traditional thinker as well as a revolutionary one, by showing how he stops short of writing symbolic formulae or solutions to problems in geometry.

The second part is much longer, and will take us through some proofs, none terribly difficult, I hope. The mathematical climax is the analysis of the inscription of a regular heptagon in a circle, an example of the exegetical art in action taken from Viète’s work.

The Importance of Viète

According to Jacob Klein, the man chiefly responsible for interest in Viète in the last 70 years, “The very nature of man’s understanding of the world is henceforth, (that is to say, after Viète’s work), governed by the symbolic number concept. In Viète’s ‘general analytic’ this symbolic concept of number appears for the first time, namely in the form of the *species*.” Now this last remark needs clarification, particularly in its use of the term “species,” that is “form” or “*eidos*,” but also, perhaps, in the notion of symbolic mathematics that Mr.

Richard Ferrier teaches at Thomas Aquinas College. He has written, but not yet published, a study of Viète’s transformative work on the analytic art.

Klein is drawing upon (*Greek Mathematical Thought and the Origin of Algebra*, p. 185). I will give a very brief account of what I think he means.

It comes to something like this. The letters upon which we operate when we do algebra are not signs of the same order as the words of common speech. They do not *immediately* intend or signify anything. They are ciphers related to each other by rules of connection analogous to syntax in a spoken language. It is by those rules alone that they acquire what “meaning” they have. That meaning is now called “syntactic” as opposed to “semantic” meaning. The equals symbol in an equation does not mean, “is the same in magnitude or number.” It is rather a relation between two other terms, say A and B, which happens to be a convertible relation, so if the first relation holds, $A = B$, so also the second, $B = A$ holds. If you give this symbol enough of the behaviors or of the characteristics that ordinary equality has, then what will be *valid* for its use, that is, what will be consistent to say of it once you have posited the suitable rules governing its role in a system of such symbols, will also be *true* of arithmetic or geometry. This presumes, of course, that you interpret all the symbols as the kind of beings that arithmetic considers, relations of inequality and equality, numbers themselves, and so forth. But prior to this interpretation, the set of symbols does not signify numbers, magnitudes, equalities, additions, divisions, or any other determinate kind of mathematical being. Moreover, the indetermination in the single letter symbols gives rise to the notion of a variable, which in turn underlies modern mathematical or formal logic. Such mathematics is symbolic mathematics. Mr. Klein claims that the symbolic concept of number originates in the work of François Viète, and that this concept comes to dominate modern mathematical thought and to influence modern conceptuality altogether.

Now I do not intend to demonstrate these larger claims or to argue for them, but rather to look into some particulars connected with Klein’s view of Viète. First let us note this

aspect of his view: Klein says, “The symbolic concept of number appears for the first time . . . in the form of the species.” How can that be? How is it that the term “species” can be the entry point for the modern notion of number, and in general for a mode of cognition, symbolic cognition, to enter mathematics?

This is what I understand Mr. Klein to mean, and I think he is correct: Viète found a use of the word “*eidos*” in a text of solutions to number problems, an ancient text, the *Arithmetica* of Diophantus, where the word is used for the unknown *number*. That is, Diophantus calls the unknown a “species.” Viète conjoined this use to a discussion of a method of finding unknown magnitudes in *geometry*. This method is called analysis. In geometrical analysis, authors such as Archimedes and Apollonius operate on unknown or not actually given lines and figures as though they were given. Viète, then, brought these two together: (1) the name “species” or “*eidos*” for the unknown or, as we would say, variable; and (2) operating or calculating without making any distinction between unknown and known quantities. But he went further. He replaced the given quantities, numbers, or magnitudes with more species, and so produced the first modern literal or non-numerical algebra, which he called “species logistic,” that is, computation in forms.

A very elementary example may serve to clarify this idea somewhat. You are familiar with the problem of finding a fourth proportional. Now if you express that in algebra you would come up with something like this: let ‘x’ stand for the unknown. x is to a as b is to c . In that expression the letter ‘x’ does not stand for something you now have or know. It is not clear how you should add or multiply or in any way deal with something that you do not yet have.

Let us ignore this mystery of operating on what is not there for you, what is not given. Writing x is to a as b is to c , you take the product of means and extremes, and you wind up with x times $c = a$ times b . Then you divide both products by c and you have $x = a$ times b divided by c , where a , b , and

c are the givens that will determine x, and when you have that you have a *formula* for making x be given to you. Of course, you cannot compute x unless a, b, and c are actual numbers, say 4, 5, and 10. If these are the given numbers, then x is given as well, and it is the number 2. Neither can you construct x, unless you have determinate magnitudes, say straight lines, as the three givens.

If you keep the letters a, b, and c and you allow a, b, and c to be equally indeterminate with x, then you no longer have a definite x so much as a kind of relation between a, b, c, and x. In algebra, we stop with this relation, calling it both the formula and the solution to our problem. That is what happens when the particular numbers of an ordinary problem—to find the fourth proportional—are replaced by signs for the possibility of finding such a number. You no longer can take a times b and divide it by c, because a, b, and c are just as indeterminate as the original unknown. This is what I mean when I say that Viète used Diophantus' term species to designate not only the unknowns but also the knowns. He called the resulting field of calculations and operations, and the consequences of performing them according to laws set out, "species logistic," that is, calculation in species.

If we take two easy steps beyond what Viète did, writing a formula for quadratics and using two unknowns as axes in a plane, his work leads directly to negative, irrational and complex numbers, to the idea of a variable, and to analytic geometry. All this stems from letting the symbol, the "species," stand in for the determinate number or magnitude of classical mathematics.

Viète himself puts it this way, (he uses the term "zetetic," which I will discuss later). "The zetetic art does not employ its logic on numbers, which was the tediousness of ancient analysts, but uses its logic through a logistic which in a new way has to do with species. This logistic is much more successful and powerful than the numerical one." So Viète commands our attention as the originator of the modern algebraic number concept, and the notion of mathematics as a symbolic

science not immediately about anything but the arrangements of its own terms into systems. I say Viète commands our attention as the originator. Now it is of course the case that however revolutionary Viète thought he was and however proud he was of his accomplishment, he did not see all the things that came from it, for example, imaginary numbers, or co-ordinate geometry. How did *he* think of his work? He considered his work largely a restoration of a way of seeking that went back to Plato. I say he considered it *largely* a restoration because there is an addition to the restored analysis that Viète expressly claims as his own, which he calls exegetic, and it is this part that I would like to set out next.

The principal ancient text from which Viète formed his view of analysis is the Seventh book of the *Collectio* of Pappus. Pappus distinguishes between seeking, or "zetetic," and providing, or "poristic" analyses. This difference answers to the difference between the propositions in the *Elements* that end "Q.E.D" and those that end "Q.E.F.," that is, between theorems and problems. There are two types of analysis, since there are two types of propositions with which geometers are ordinarily concerned. Viète, for reasons too subtle to lay out here, read Pappus in another way, making zetetic and poristic parts of one procedure, a procedure ordinarily applied to a problem, not a theorem. Still, he conceived of his work as a restoration of the lost art—or perhaps the partially lost art—of analysis, that is, as a renovation, not an innovation. That is a striking characteristic of Viète and his contemporaries. You find in them a preference for the lost or partially recovered sources in antiquity, for Democritus over Aristotle, and when Plato is not fully available, for Plato over Aristotle, for Archimedes over Euclid, and in general for whatever the schoolmen did not hand on over what the schoolmen did hand on. They sought the key to the deepest truths in the arts and sciences in a correct restoration of the faultily preserved sources of antiquity in preference to accepting the ones that they had received more perfectly intact from their teachers.

This is what Viète says: “And although the ancients had set forth a two-fold analysis, the zetetic and the poristic, it is nevertheless fitting that there be established also a third kind, which may be called ‘rhetic’, [telling] or ‘exegetic’, [showing or exhibiting], so that there is a zetetic art by which is found the equation or proportion between the magnitude that is being sought, and those that are given, a poristic art, by which from the equation or proportion the truth of the theorem set up is examined, and an exegetic art, by which from the equation set up or the proportion there is exhibited the magnitude itself which is being sought.” As if to emphasize his own achievement in completing the method of analysis, he says, “And thus the whole threefold analytical art, claiming for itself this office may be defined as the science of right finding in mathematics.” As to the importance of the third part, rhetic or exegetic, he says in another place, “Rhetic and exegetic must be considered to be most powerfully pertinent to the establishment of the art, since the two remaining provide examples rather than rules.” He emphasizes the importance of the third part in the remarkable concluding words of the introduction to the analytical art: “Finally, the analytical art, having at last been put into the threefold form of zetetic, poristic, and exegetic, appropriates to itself by rights the proud problem of problems: To leave no problem unsolved.”

What exactly is the third part of the analytic art, the showing or exegetic part, the part that Viète himself emphasized and considered his own invention? One way to proceed to answer would be to read or re-read the parts of Viète’s *Isagoge* that touch on exegetics. If your experience is like mine, though, you know that it is one thing to have someone describe a procedure, especially a mathematical procedure, another to know it from having carried it out. Accordingly, I propose to explicate Viète’s notion of exegesis by doing some geometry. To prepare for the example of exegetics taken from Viète, we will look at a problem of my own.

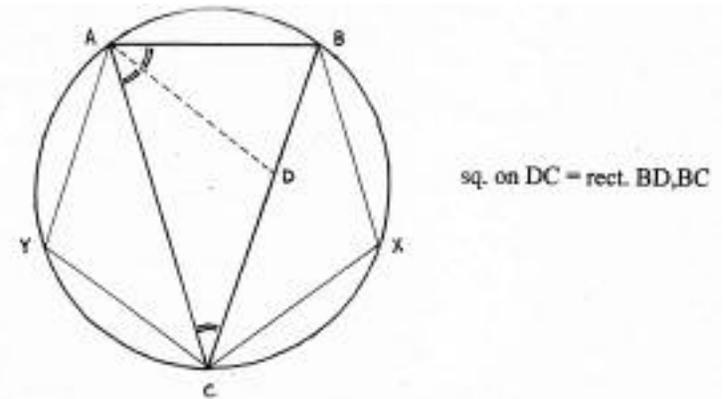
Viète’s Exegetic Art

Analysis Bootcamp

Our practice will be the analysis of the construction of a regular polygon, the pentagon.

Problem: to inscribe a regular pentagon in a circle.

Figure 1 *Inscription of a regular pentagon in a circle*



Now the analysis of a construction always starts this way: suppose you already have what you want to make.

Let it be done, or rather as they say, “let it have been done.” Our figure is, then, given not really but hypothetically. ABXCY is supposed to be a regular pentagon in a circle. All the sides are equal, equal sides subtend equal arcs, equal arcs subtend equal angles at the circumference, and therefore angle ACB is precisely half the angle ABC, and is also half the angle CAB.

The triangle CAB is therefore an isosceles triangle, CA equaling BC, with the vertex angle half the angles at the base.

Next let the angle at A be bisected by AD. The two half angles then will each be equal to the angle ACD and the triangle ADC will also be an isosceles triangle.

Since the angle BAD is equal to the angle ACD, if a circle be described through the three points A, D, C, the line AB will be tangent to that circle, for the angle that a tangent makes with a chord is equal to the angle subtended by the same arc at the circumference. But the square on a tangent is equal to the rectangle contained by the whole secant and the part outside the circle, that is BC and BD, i.e., the square on AB = rectangle BD, BC. But inasmuch as AB = AD = DC, then the square on DC = the rectangle BD, BC.

Well, then, if I had this pentagon in this circle, I would have this line BC cut at a point D so that the square on DC, the greater segment, is equal to the rectangle contained by the whole line and the lesser segment. But there is a proposition that tells me how to do that, the eleventh proposition in the second book of the *Elements*. So, let it be done and then prove “forwards,” or synthetically, all the connections that I just established “backwards,” or analytically, and at the end we will really have our pentagon in a circle.

Now, that is a classical geometrical analysis and synthesis of the sort you would find in the second book of Apollonius, numerous places in Archimedes, and in Pappus. I have not used anything algebraic. But, let us indulge ourselves in some algebra, and look at the equation here: the square on DC = BD, BC. Calling BC ‘a’, and DC ‘x’, the equation is $x^2 = a(a-x)$. “And,” you could say to yourself, “if I could only solve for x, then I would know how long to make DC, and if I could make DC then I could do the problem. But this is a quadratic equation, and I can solve it.”

That is the method of analysis. I include here what came to be called the ‘resolution.’ We began in the classical mode by looking at the angles and finding a proportion, which we resolved into an equality between a square and a rectangle. When we reach this point, we can either continue in the classical mode by thinking of the equality in geometrical terms, and do the synthesis as a construction, which would be the answer to the problem. Or we can examine the equality alge-

braically and work with it in that mode to reach a formula for x, and consider the formula the solution of the problem.

Recall what Viète said about exegetics: it is supposed to provide the unknown magnitude itself, to exhibit or construct it. One would expect exegetics to provide the way for finding such an unknown as we had in the present problem, that is, it would seem to be some sort of procedure for cutting the line BD at a point like D that will produce the requisite properties. One is tempted to think that the procedure intended is to use the quadratic formula to solve for x, and then interpret the right hand side as a series of simple geometric constructions. As we shall see later, though, this is an error. Viète’s exegetic art is not the mechanical exploitation of the formulae of solved equations.

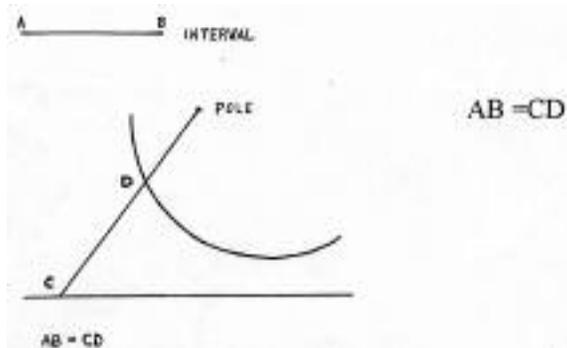
Analysis with Live Ammunition

Next, we will see what Viète does in a real problem that is considerably more complex but of greater interest.

The problem is to inscribe a regular seven-sided polygon, a heptagon, in a circle. It is taken from Viète’s published work and is a real instance of the kind of solution Viète promises in the introduction, that is, it is a real instance, I think, of exegetics. To solve it we will need a postulate and two lemmas.

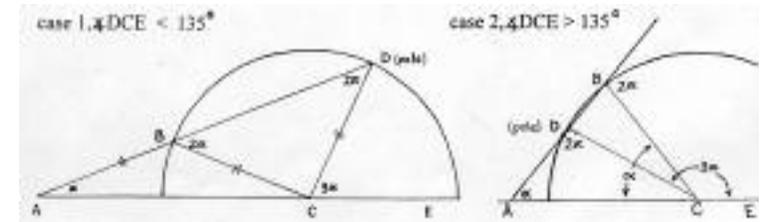
Postulate: To draw a straight line from any point to any two lines so that they cut off on it any possible predetermined interval. (I have illustrated this in the next figure. See figure 2.)

Figure 2 Postulate



The point called the pole is the “any point” of the postulate. “To draw a straight line from any point”: so the straight line must pass through the pole. “To draw it to any two lines”—those two lines are the curved line and the straight line below—“so that they cut off on it any possible predetermined interval”—the interval is up there above; it is given in advance, predetermined, and what I would like you to grant me in this postulate is the power to put a line through that pole so that between those two other lines, that interval is cut off. Now I must say, “Any *possible* predetermined interval” because as is obvious to the eye, I guess, in this figure, if I made that interval too short, I could not fit the line through the pole so as to have that interval cut off on it no matter what. If the two given lines were, for another example, concentric circles, and if the pole were the center of the circle, then this postulate would be of no use at all because the only distance between those two lines on lines drawn from the center of the circle is the difference between their radii; that is all I could have. So I must say, “any possible predetermined interval.” If it is possible, I would like the right to place that line passing through the pole so as to have that distance intercepted on it. This postulate is mentioned in the *Isagoge*. To the reader who has not read more of Viète than the *Isagoge*, it is not clear why that is in there at all—but he asks for that postulate for certain purposes, and I ask for it, too.

Figure 3 First Lemma: Trisection of an angle:



First Lemma: To trisect a given angle.

Let the given angle be DCE.

Let a circle be described with C as center and radius CD.

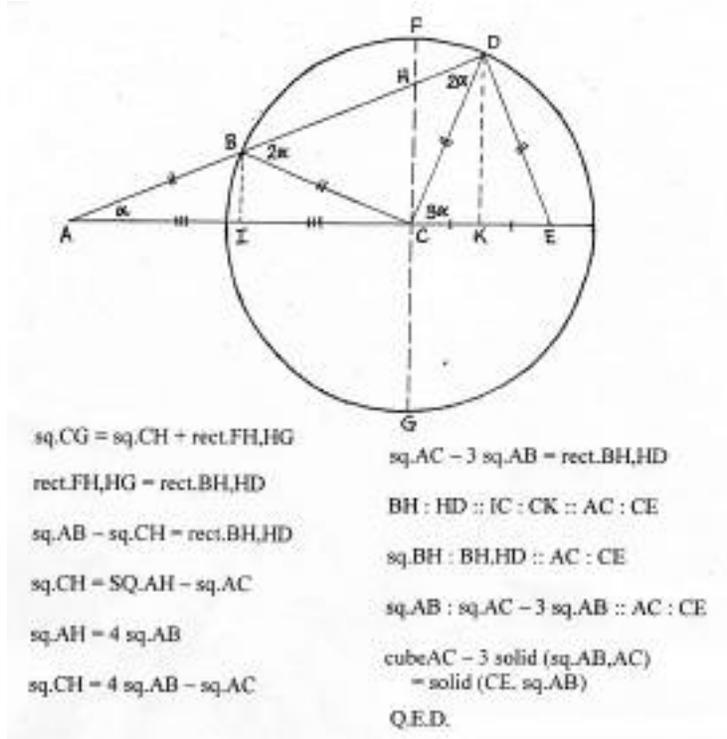
With pole D and interval = CD we use the postulate to insert line ABD, so that the interval falls between the circle and the line CE extended.

I say that angle BAC is $1/3$ of angle DCE.

Since $AB = BC = CD$, the triangle BCD is isosceles, with equal angles at B and D. But each of these is double the angle at A. And the one at D together with A is equal to angle DCE. In the case where the angle to be trisected is greater than 135 degrees, we use the equal exterior angles at B and D, which are again each twice the angle at A, as in the second figure. Q.E.F.

Second Lemma: In the first figure for Lemma 1, if we construct $DE = CD$, the following equality holds:

The cube on AC, minus three times the solid contained by AC and the square on AB, is equal to the solid contained by CE and the square on AB.

Figure 4 Lemma 2: $\text{cube } AC - 3 \text{ times solid } (AC, \text{sq. } AB) = \text{solid } (CE, \text{sq. } AB)$ 

Looking at the figure (figure 4), if I drop perpendiculars BI, DK, and erect a perpendicular CH and extend it, on the vertical line FHC_G I have the situation presented in the Elements 2.5, that is, I have a straight line bisected and cut somewhere else, and therefore the square on the half will equal the sum of the rectangle contained by the unequal segments and the square on the segment between the bisection point and the point making the unequal cut. And looking at the straight lines BHD and FHG, I have the situation of 3.35, namely two lines in a circle cutting each other. When two lines in a circle cut each other, the rectangles contained by the segments are equal. I will use those properties and the Pythagorean Theorem and a few other elementary truths to

demonstrate the lemma, but the key ones are the ones I just indicated.

The proof is given next to the figure. The square on AB, which, since $AB = BC$ is the same as the square on CG, is equal by that proposition in the second book of the *Elements* to the square on CH and the rectangle contained by the unequal segments FH, HG. But the square on AB minus the square on CH will be the rectangle which is in turn equal to the rectangle BH, HD by the other proposition from the third book, namely that rectangles made of the segments of two intersecting lines in a circle are equal. Now the square on CH, by the Pythagorean Theorem, will be equal to the difference between the square on the hypotenuse AH and the square on AC. But since BI is a perpendicular drawn from the vertex of an isosceles triangle, it will bisect the base AC, and, BI being parallel to HC, the line AH will also be bisected at B, so that the square on AH is four times the square on its half, AB. Then the square on CH is the difference between four times the square on AB and the square on AC. Next, from that and the line above, namely, that the difference between the square on AB and the square on CH is equal to the rectangle BH HD, the square on AC minus three times the square on AB will be equal to the rectangle BH HD. Now that is almost where I want to be because I am interested in the cube on AC, and the difference between that and—I am looking at the conclusion now—three times the solid contained by AC and the square on AB. I have those solids with the height removed, as it were, I just need the height AC and I have those solids. That is my next goal.

Then BH is to HD as IC is to CK, since any two lines cut by parallel lines are cut proportionally. And if BH is to HD as IC is to CK, then it is also in the same ratio as their doubles AC and CE. So that BH is to HD as AC is to CE. Then, taking the common height, BH, the square on BH is to the rectangle BH, HD as AC is to CE. But BH is equal to AB so that the square on AB is to the difference between the square on AC and three times the square on AB—that was what we

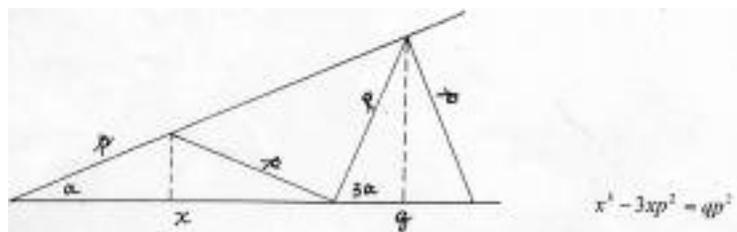
found equal to the rectangle BH, HD—in the same ratio, namely as AC is to CE.

Then taking means and extremes, the solid contained by the means is equal to the solid contained by the extremes, namely the cube on AC, minus three times the solid contained by AC and the square on AB is equal to the solid contained by CE and the square on AB. Q.E.D.

I would note at this point that everything we have gone through in these preliminaries is strictly classical, synthetic geometry. It could all have been done by Archimedes or Apollonius, and some of it, in fact, was done by them.

Now let us see what interest this second lemma might hold. It seems to me the easiest thing would be to look at this figure given below (figure 5) which removes all the middle steps and just looks at the result.

Figure 5



If you have two isosceles triangles with equal sides p and unequal sides q and x , then you get an equation answering to the equality that I gave in the geometrical mode namely that $x^3 - 3xp^2 = qp^2$ or, in short, this is a geometrical configuration that answers to a cubic equation. If you have ever gone back and looked at things in Euclid, trying to express them in equations, you will know that cubics are unusual. You almost always get squares. This theorem answers to a cubic equation. And that is why Viète proves it. He proves this theorem in order to give a geometrical counterpart to that cubic equation.

If the terms in the equations are interpreted as the sides and bases of the triangles in this theorem, then the solution

of the interpreted equation would reduce to the construction of this figure. That is, if you could construct the triangles set up this way, and in particular if you could give yourself x , then you would have solved this equation, this interpreted equation, as I would like to speak of it. To put this another way, if $x^3 - 3xp^2 + qp^2$, then x can be found as the base of an isosceles triangle whose equal sides are p and whose base angle is $1/3$ the base angle of an isosceles triangle with sides p and base q . Now the angles of such a triangle, namely the one with two sides p and a base q , are given. Why? Because the whole triangle is given. If you have three sides you have a determinate triangle, and so the angles are given. And given two equal sides and any one angle in an isosceles triangle, the remaining side and the remaining angles are also given. Consequently the base of the flatter triangle would be given. Thus all we need to do to solve for x is to trisect a given angle. This is easily done by means of our first lemma.

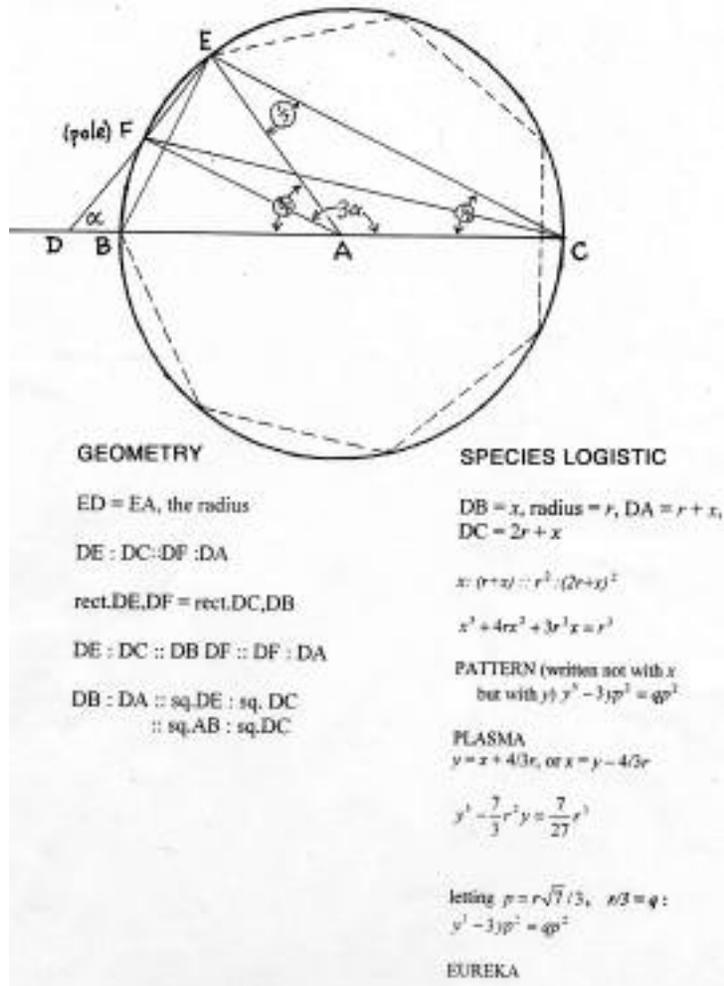
To review: we have a postulate that allows us to insert a rotating line through a given point across two lines so as to have cut off on it any interval. We have found a figure that answers to a certain cubic equation, namely, paired isosceles triangles with a 3:1 base angle ratio, that answers to a certain cubic equation, and we obtained this figure by trisecting an angle. Those are the preliminaries.

Last, let us see how Viète inscribes a regular seven-sided polygon in a circle. This problem, incidentally, is Viète's own final problem in his *Supplementum Geometriae*, his completion of geometry. I have chosen it not only for its intrinsic interest but also because I think Viète, who calls the *Supplementum* a work in exegetics, gave this problem as a specimen of the workings of his exegetic art.

Problem: To inscribe a regular heptagon in a circle.

Let it have been done. (See figure 6.)

Figure 6 Inscrition of a regular heptagon in a circle



And let BE be one side of the heptagon; then the angle ECB will be 1/7 of two right angles. Then the angle at E, CEA, will also be 1/7 because that is an isosceles triangle, EAC. Then the angle EAB will be 2/7 because it is the sum of those. Now from the point E let ED be drawn equal to the radius of the circle, and where it cuts the circle at F let the line FA be

drawn to the center. Then we have two isosceles triangles of the sort in the trisection lemma from which it follows that the angle FAC is three times angle EDA. It is the same figure that we had in the trisection lemma. In the triangle EAC the two smaller angles each being 1/7 of two right angles, the remaining angle at A is 5/7 of two right angles, but the whole angle FAC being triple the angle at D, which is itself 2/7, will be 6/7. The remaining angle FAE will be 1/7 of two right angles. It will be equal then to the angle AEC, so that the lines FA and EC will be parallel.

Since those two lines are parallel they cut the sides of the triangles DEC and DFA proportionately, so DE is to DC as DF is to DA. But, by another proposition from book 3 of the *Elements*, the book of circles, the rectangle contained by the segments of a secant, that is, the whole secant and the part outside, DF, DE, is equal to the rectangle contained by the segments of any other secant drawn from the same point outside. So DB, DC equals DE, DF. Then, turning that equality into a proportion: DE is to DC as DB is to DF, but as DE is to DC as DF is to DA, therefore DB, DF, DA are continuously proportional. Then, in a continuous proportion the first is to the third in the duplicate ratio, or as the squares, on any terms in the ratios and also any terms having the same ratio as those terms. So DB is to DA as square DE is to square DC. But DE was by construction made equal to the radius so DB is to DA as square AB is to square DC. Here we stop.

How does one know when to stop in an analysis? Well, here is one way I know in this one. You stop when you have projected the ratios you know in various parts of the figure down to ratios on one line. It happens in the synthetic proofs of Apollonius, too. You take a number of ratios in the figure and when you transform them into a proportion all on a single line, that tells you how the lengths in that line are related. That is where we are now. We have taken the ratios that exist through the figure because of its shape and we have turned them into a proportion among the parts of one line. That is also beautifully adapted for conversion into an equation.

That line, of course, is the diameter extended, DBAC, and our proportion is DB is to DA as square AB is to square DC.

Let us turn this into species logistic, or algebra. What do we need to construct the heptagon that we do not know? Is it not clear that we need to know DB? If we knew DB, we could locate point D and since DE is equal to a radius, we could just swing a radius out from D and cut off the other end of the side of the heptagon, and we would be done. So DB is the unknown, and that is the only term we need to find on this line, on the diameter extended. So call DB 'x', and the radius 'r'. DA will then be (r+x), and DC will be (2r+x). The proportion then may be written: $x : (r+x) :: r : (2r+x)$. Means and extremes can be taken, yielding the equation: $x^3 + 4rx^2 + 3r^2x = r^3$.

This does not look very promising. It is a cubic, and we have a pattern for solving a cubic, but not this cubic. Happily, there is an algebraic gimmick that Viète invented called "plasma," which will do the trick. (Almost none of Viète's technical terms passed into common use, by the way. Accordingly, it is like reading ancient law-books to read him talking about his procedure. One of the few Viètean terms that did last was "coefficient," but "plasma" did not.) "Plasma" is the technique of getting rid of an unwanted term in an algebraic expression by taking a substitute variable. In a way you might say that is what you do when you complete the square in solving a quadratic. You find the right thing to remove the middle term, the term that is just an x, and then it is just a simple square; it is no longer a three-term expression. Now in cubics if you could remove two terms at once, then all you would have to do is take cube roots and you can solve them in much the same way you solve quadratics. Plasma gets rid of one term, a term in x or x squared, but unfortunately only one. The substitution that you make to do it is $y = x + 4/3r$, or $x = y - 4/3r$. Let that substitution have been made, and you get the equation in y: $y^3 - \frac{7}{3}r^2y = \frac{7}{27}r^3$.

Now if we let $p = r \sqrt{7/3}$, and $r/3$ be q, this equation transforms into our pattern $y^3 - 3yp^2 = qp^2$. Point D gives x and y, and hence one side of our heptagon. Since we have the side, we can just go around the circle six more times and the problem is done.

I omit a slight digression in Viète's analysis here. It is done for reasons of elegance and does not substantially change the argument. This is the pivot point of the argument, where we move from analysis to synthesis. Viète's synthesis runs through the analysis backwards: first the analysis in the equations and proportions, and then the analysis in the angles, until he has a proof that the figure that he had made in the circle is a regular seven-sided figure. I will omit the synthesis.

It must be noted, however, that not only would every ancient geometer have provided a synthesis, but that Viète, too, gives it in the *Supplementum Geometria*, the text from which this all comes. Indeed, it is *all* that he gives. What I have laid out here is a reconstructed analysis following his remarks on how to do these sorts of problems. That is, he says that the proof will be the reverse of the analysis and that the analyst dissimulates and does not show you everything he did. This seems to me to be a kind of challenge to readers of those texts to find out what he did do. Descartes says just that in his *Géométrie*. So that is what I have given you: it is a reconstructed analysis of his problem. Like Descartes, Viète complains that the Ancients covered their tracks and then does the same himself!

Conclusion

A survey of our procedure is now in order. First, we did not do something that the discussion of exegetics in the *Isagoge* might suggest. I think also it is what Mr. Klein's book suggests, and I know it is what I thought when I first read the *Isagoge* myself. We did not, using algebra, or as Viète calls it, species logistic, solve the equation to which we had reduced our problem. That equation would be the first one that had

the fours in it, and that we later on reduced by plasma to another equation. At no point in working with that equation did we solve for the unknown in the shape of a formula. There is not an algebraic or analytic solution to that equation; it is not to be found. You perhaps may wonder if I misspoke my thought there: absolutely not. In fact, the cubic equation involved here is the so-called “irreducible cubic,” which means that it cannot be solved by algebra. If you try to solve it by algebra, you involve yourself with imaginary quantities, and every attempt you make to get rid of them is like stepping on the bump in the rug: they pop up again elsewhere. They cannot be gotten rid of. So it is not just that he failed to do it. He cannot do it; no one can do it.

Exegetics, then, does not mean performing the geometric counterpart of an algebraic solution *as it is laid out in a formula*. It does not mean getting $\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ for example, and then going through the squaring, adding, subtracting, and so on in geometry, with the formula laid out for you as a kind of blueprint. This is something to which the opening pages of Descartes' *Géométrie* point. I say “point” because I am not sure even Descartes does it; you have to think for awhile to see whether he intends his figures to be the simple exploitation of his operation or not. Perhaps they are not exploitations of the quadratic formula, but independently given geometrical constructions. As is so often the case with the foxy Descartes, he does not say enough to help us be sure what he has in mind. It is defensible, though, in the light of what he *does* say, that he intends you to have a geometric counterpart for each algebraic operation, and just to exploit the formula directly. That is at least defensible in Descartes. In Viète it is not so.

So what did we do if we did not do that? Well, first we performed an ordinary geometrical analysis of the sort Archimedes might have done. This ended in a proportion. That proportion was then treated as an item in logistic. We forgot that it was about lines and just regarded it as a number of items in a logistic relation. It was transformed until it

reached a certain form. What form? A form we knew in advance from our first lemma to have its geometrical counterpart in a constructable figure—that is, constructable if you grant our postulate. Then, after we had noted certain proper elegances pertaining to this particular problem, we proceeded synthetically (or rather, we would have if we had gone through all of the steps), first constructing the requisite line and then the whole polygon, and then, via the resulting equation and other relations, proving that it is indeed a regular seven-sided figure in a given circle. Zetetic is Viète's name for the first two parts of this procedure, both the geometrical part terminating in the proportion and the algebraic analysis in which we reach an equation in standard form.

Zetetic ended when the equation to which the problem had been reduced fit one of a number of standard forms, our lemma for cubics being one such form. This lemma provides the principal exegetical part. By means of this lemma, we can now find the unknown, in this case the line DB. Exegetics as a procedure is regular and synthetic. I mean those terms strictly, that is, the construction is a standard, mechanical exploitation of the equation that can be given by a standard construction. The complete synthesis simply runs through the zetetics in reverse order. If your problem reduces to a cubic of another form (this is the same with the one we have been examining, but with the signs on the left side of the equation switched: $3p^2x - x^3 = p^2q$) there is, in the *Supplementum*, another, similar exegetical lemma you can use to start your construction, and the situation is quite the same for problems such as inscribing the pentagon, reducing to quadratic equations.

Viète treats quadratics in another treatise, and it is an extraordinarily puzzling little work. Almost everything in it is painfully obvious, at least the first 14 or so propositions are, and it has the rather odd and stuffy title, which I think should now make sense, “The Standard Enumeration of Geometrical Results.” It is a handbook; it is like a carpenter's manual. When you have one of these, build in this order that, etc. He

does this for quadratic equations and for certain biquadratic equations. Moreover, in all these cases, the proof, which follows the construction, as in Euclid, is a straightforward reversal of the geometrical and logistic analyses that preceded—now that, I think, fits Viète's language in his introduction; it makes sense of his language.

So exegetics in practice comes to this. It is the provision and employment of a series of standard constructions for finding unknown quantities when such quantities are enmeshed in equations of various degrees, *without* solving the equations. This is important, since it manifests the way Viète stands in two camps, modern or symbolic, and ancient or constructive. Though species logistic is of profound value to him, he never rests in the fully analytical or symbolic solution to his equations. They must always be perfected by realization in construction or computation, artful synthetic and non-symbolic procedures.

I hope we may now see why Viète wrote, "Exegetic comprises a series of rules, and is therefore to be considered the most important part of analytic, for these rules first confer on the analytic art its character as an art, while zetetic and poristic consist essentially of examples." The analyst who knows exegetics will know what the standard constructable forms in each degree are, and he will accordingly know what to aim at in his zetesis, in his analysis. This will in turn inform his development of the techniques of species logistic, and finally, if he believes, as Viète came to believe, that he has all the exegetics that can be, he may well boast that, "The analytical art, having at last been put into the threefold form of zetetic, poristic, and exegetic, appropriates to itself by right the proud problem of problems, which is to leave no problem unsolved."