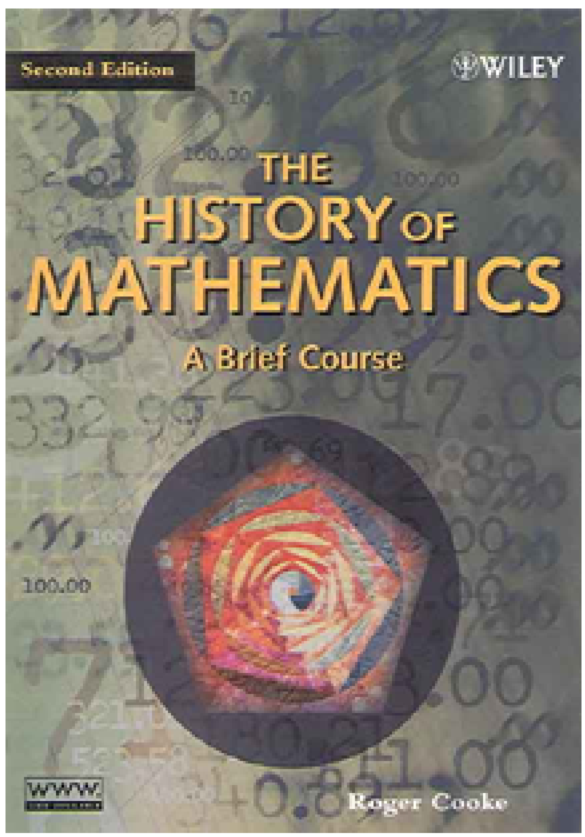
The History of Mathematics: A Brief Course

Answers to Questions and Problems

Roger Cooke
University of Vermont



Contents

Preface

This “solutions manual” is not intended for students, although I leave it to the judgment of the instructor whether to post part or all of it online for student use. Most of the speculative questions that I ask have so many possible “correct” answers that it is nearly pointless to suggest possible answers. The suggested answers in this book are off-the-top-of-my-head responses to questions that I thought of months ago when I was writing the book itself. In many cases they are not the answers that I had in mind when I originally wrote the question. In many more cases, they are merely my own idiosyncratic views about the kind of world I would personally like to live in, a world that others may dislike. They are seldom the ripe fruit of many years of reflection, and if they give the instructor the chance to say that the author is a fool and deliver a withering refutation of them, that also helps promote the exchange of ideas and provokes students to think for themselves. In short, I have rather “let myself go” in writing these answers, in a way that I would not do when attempting to impart information. The answers are replete with my personal *obiter dicta* on a variety of social and philosophical questions on which I have no particular expertise. I have spent many decades reading the works of the great philosophers, but have never taken a course in philosophy. Thus I have a lot of ideas of my own, utterly untested against the sophistication of a well-informed philosopher. Most undergraduate students are in the same situation, a few of them being a step ahead of me on this score. Hence we meet on more or less equal grounds. The result, I hope, will be at least interesting, if occasionally infuriating.

Since the problems are taken from the book itself, I have left all literature citations in these solutions just as in the text, but I have not included the list of literature from the text. Likewise, there are references to figures and equations in these solutions, and those will make sense only if you have a copy of the book at hand as you read.

Roger Cooke
March 2005

CHAPTER 1

The Origin and Prehistory of Mathematics

1.1. At what point do you find it necessary to count in order to say how large a collection is? Can you look at a word such as *tendentious* and see immediately how many letters it has? The American writer Henry Thoreau (1817–1863) was said to have the ability to pick up exactly one dozen pencils out of a pile. Try as an experiment to determine the largest number of pencils you can pick up out of a pile without counting. The point of this exercise is to see where direct perception needs to be replaced by counting.

Answer. Answers to this question will vary from person to person. With some practice (and much experience working cryptic crossword puzzles), I have learned to group words into syllables at a glance and sum the numbers of letters. That is not the same thing as direct perception of the number. Generally symmetry or asymmetry distinguishes words with an even or odd number of letters. Up to about 8 letters, direct perception is possible. I am not able to come anywhere close to Thoreau's legendary ability. Beyond five or six, the hand becomes very unreliable.

1.2. In what practical contexts of everyday life are the fundamental operations of arithmetic—addition, subtraction, multiplication, and division—needed? Give at least two examples of the use of each. How do these operations apply to the problems for which the theory of proportion was invented?

Answer. Addition is needed if you keep a running total of what you are spending while you shop and if you are "counting calories" on a diet. Subtraction is useful in entering checks into the register of a checkbook or figuring out how much change you have due when making a purchase with a large bill. Multiplication enables you to compute how much food to buy when planning a party (you multiply the number of guests by the allotted consumption per person) and how much money you have earned, when you are paid by piecework or by the hour. Division is used to compute the fuel economy of an automobile. (Usually, the distance traveled is divided by the amount of fuel consumed to yield the number of kilometers per liter or, in the USA, miles per gallon.) It is also important wherever sports records such as earned-run averages and batting averages are kept.

Proportion depends on the operations of division and multiplication. The proportion $A : B :: C : D$ is expressed numerically by measuring A, B, C, and D to get numbers a, b, c, and d respectively. The proportion is then interpreted as the numerical equality $\frac{a}{b} = \frac{c}{d}$ or, what is the same, equality of the products $ad = bc$.

1.3. What significance might there be in the fact that there are three columns of notches on the Ishango Bone? What might be the significance of the numbers of notches in the three series?

Answer. Just to recall what is in the text, one column contains the series 11, 21, 19, 9; the second contains 11, 13, 17, and 19; and the third contains 3, 6, 4, 8, 10, 5, 5, and

7. If the bone was used for a practical purpose, the three columns could correspond to physical *locations* and the successive numbers in each column to successive time periods, in which certain objects of interest were counted at the locations. Or the roles of space and time might be reversed here. Since the potential numbers of objects, places, and times is so large, we have virtually endless possibilities with this interpretation of the usefulness of the bone, ranging from peaceful gathering of eggs or berries at one extreme to deadly warfare at the other.

If the purpose was purely esthetic or intellectual, it is striking that so many odd numbers occur, and one should try to explain that fact. Were these numbers the lengths of parts of a geometric figure? Or perhaps the lengths of strings on a musical instrument? Again, the possibilities are endless, and it does not seem likely that anyone today can narrow them down so that some will be noticeably more probable than others.

1.4. Is it possible that the Ishango Bone was used for divination? Can you think of a way in which it could be used for this purpose?

Answer. As the discussion of divination in Chapter 7 shows, one needs some physical device that can display various outcomes. Unless the Ishango Bone was part of a set, it probably was not used for divination. It would have to be used in some way so that one of the 16 numbers on it could be distinguished in some "random trial," like the top face on a die.

1.5. Is it significant that one of the yarrow sticks is isolated at the beginning of each step in the Chinese divination procedure described above? What difference does this step make in the outcome?

Answer. If this step were not taken, the number of different possible outcomes would be greatly reduced. This extra step enables the "randomness" of modular arithmetic to exert its effect. (That randomness is what makes modular arithmetic the basis of RSA codes.)

1.6. Measuring a continuous object involves finding its ratio to some standard unit. For example, when you measure out one-third of a cup of flour in a recipe, you are choosing a quantity of flour whose ratio to the standard cup is $1:3$. Suppose that you have a standard cup without calibrations, a second cup of unknown size, and a large bowl. How could you determine the volume of the second cup?

Answer. You could count the number of standard cups needed to fill the large bowl with water, then count the number of cupfuls of the second cup needed to fill the same bowl. If the bowl is very large, the two numbers will be approximately in the inverse ratio of the two volumes.

1.7. Units of time, such as a day, a month, and a year, have ratios. In fact you probably know that a year is about $365\frac{1}{4}$ days long. Imagine that you had never been taught that fact. How would you—how did people originally—determine how many days there are in a year?

Answer. You could put a stake in the ground and each morning at sunrise place another stake some reasonable distance to the east of the first stake, right in line with the rising sun. That second stake will move back and forth from south to north (if you are not too close to the equator). If you count the number of days between its extreme points (solstices) over a period of many years, you will get an equation of the form m days $\approx n$years, from which the ratio can be worked out.

1.8. Why is a calendar needed by an organized society? Would a very small society (consisting of, say, a few dozen families) require a calendar if it engaged mostly in hunting, fishing, and gathering vegetable food? What if the principal economic activity involved following a reindeer herd? What if it involved tending a herd of domestic animals? Finally, what if it involved planting and tending crops?

Answer. Any society that is dependent on growing seasons or needs to adapt to changing temperatures must be able to plan its years, knowing when to plant and when to stock up on fuel and food for the winter. Nomadic societies probably have little need of this, since the herds that they live upon control where they go. Pastoral and agricultural societies have more need, since there are breeding, planting, and harvesting seasons. In most societies these events become both social and religious occasions, and the civil and religious festivals accompanying them need to be planned. Although the underlying economic activity is tied to the sun rather than the moon, the moon is much easier to keep track of, and it is noteworthy that many religious festivals such as Easter, Passover, and Ramadan are scheduled in terms of the moon.

1.9. Describe three different ways of measuring time, based on different physical principles. Are all three ways equally applicable to all lengths of time?

Answer. One of the oldest clocks is the water clock, which presumes that the time required for a bowl with a hole in the bottom to empty is the same each time the bowl is filled with water. Another is the pendulum clock, which works on the principle that the oscillations of a pendulum are isochronous. Still a third is the movement of the stars, moon, and sun across the sky. The water clock obviously is useful for keeping track of one specific length of time. The pendulum clock can keep track of smaller units for longer times. The star clock, in conjunction with very accurate telescopes, was for a long time the standard of accuracy in time measurement, only recently replaced by the oscillations of certain crystals.

1.10. In what sense is it possible to know the *exact* value of a number such as $\sqrt{2}$? Obviously, if a number is to be known only by its decimal expansion, nobody does know and nobody ever will know the exact value of this number. What immediate practical consequences, if any, does this fact have? Is there any other sense in which one could be said to know this number *exactly*? If there are no direct consequences of being ignorant of its exact value, is there any practical value in having the *concept* of an exact square root of 2? Why not simply replace it by a suitable approximation such as 1.41421? Consider also other "irrational" numbers, such as π, e, and $\Phi = (1+\sqrt{5})/2$. What is the value of having the *concept* of such numbers as opposed to approximate rational replacements for them?

Answer. This question goes to the very heart of abstract mathematics. Of what value are *any* abstractions? The chief value is that they provide a simplified model of what we observe in the real world. If the important fact about the diagonal d of a square, for example, is that the square on it equals twice the square of which it is the diagonal, then that fact is captured by saying that the length of d is $\sqrt{2}s$, where s is the length of the side, much better and more compactly than by saying that the diagonal is "approximately $\frac{17}{12}$ of the side." It is really better to remember that $d = \sqrt{2}s$ than to choose some approximation that may vary from one situation to another, *even if the basic relation* $d^2 = 2s^2$ *is only approximate*.

1.11. Find a unicursal tracing of the graph shown in Fig. 1.

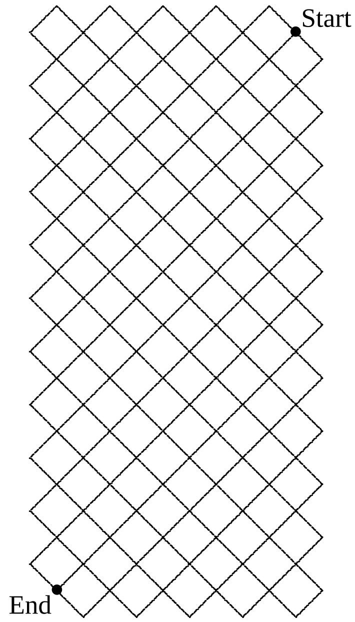


FIGURE 1. A graph for which a unicursal tracing is possible.

Answer. Imagine you are a tennis ball bouncing around.

1.12. Does the development of personal knowledge of mathematics mirror the historical development of the subject? That is, do we learn mathematical concepts as individuals in the same order in which these concepts appeared historically?

Answer. Not always. Rigorous proof in geometry which is taught (or at least, *used to be taught*) during the sophomore year in American high schools, preceded by a thousand years the algebra that is taught to middle-school and first-year high school students. And, as mentioned in connection with the work of Piaget, children learn topology intuitively long before they learn even the geometry of measurement. On the other hand, personal knowledge of arithmetic does more or less follow what seems to have been the historical order of development.

1.13. Topology, which may be unfamiliar to you, studies (among other things) the mathematical properties of knots, which have been familiar to the human race at least as long as most of the subject matter of geometry. Why was such a familiar object not studied mathematically until the twentieth century?

Answer. The secret of studying knots and other topological objects, such as manifolds, is the application of analysis and algebra, including modern algebra. One really cannot make much progress at classifying knots, except on a purely empirical level, without homotopy theory. Similarly, human behavior has been right under our noses for thousands of years, but we understand better what is happening at the center of Betelgeuse than what is happening in the center of our neighbor's brain. It's a matter of complexity.

1.14. One aspect of symbolism that has played a large role in human history is the mystical identification of things that exhibit analogous relations. The divination practiced by the

Malagasy is one example, and there are hundreds of others: astrology, alchemy, numerology, tarot cards, palm reading, and the like, down to the many odd beliefs in the effects of different foods based on their color and shape. Even if we dismiss the validity of such divination (as the author does), is there any value for science in the development of these subjects?

Answer. Here again, only personal answers can be given. As a hard-headed skeptic, I believe that mysticism has value only in the early stages of science, when intuition is the only available tool of analysis. Such intuition is usually based on some kind of analogy or proportion. This kind of informal speculation preceding what the philosopher Charles Coulston Gillispie called the *edge of objectivity*,[1] can become quite elaborate, as one can see from the cosmology in Plato's *Timaeus* and the phlogiston theory of combustion. Sooner or later, the critical faculty should be applied, at which point—I claim—paranormalists ought not to be funded out of taxpayers' money. I do not mean to assert absolutely that the paranormalists (whether they are investigating telepathy, astrology, dowsing, or other "psychic" phenomena) are wrong, although I think it highly probable that they are wrong. I grant that it is logically possible that the universe contains "sporadic" events not subject to investigation by the ordinary rules of science. But if such events can neither be controlled nor understood, they do not reach the level of definiteness that we normally associate with science, and should not be considered science. Again, I do not assert that all of reality is amenable to study by science, and I am quite content that those who consider such investigations promising should pursue them—just let them raise the funds to do so privately.

1.15. What function does logic fulfill in mathematics? Is it needed to provide a psychological feeling of confidence in a mathematical rule or assertion? Consider, for example, any simple computer program that you may have written. What really gave you confidence that it worked? Was it your logical analysis of the operations involved, or was it empirical testing on an actual computer with a large variety of different input data?

Answer. Again, a personal answer is called for. The older I become, the less confident I am about any chain of reasoning, and the more I am convinced that the main function of proof is to exhibit the logical interrelations among the parts of a theory. Its function in assuring that a given result is correct is decidedly secondary. Merely to know *that* a proposition is true, without knowing *why* it is true is of limited value. And, when I have written a program of any complexity at all, I like to test it with extreme input data to see how robust it is. Experience has taught me that I am far more likely to find the bugs that way than by pretending I am the computer and executing the operations on paper.

1.16. Logic enters the mathematics curriculum in high-school geometry. The reason for introducing it at that stage is historical: Formal treatises with axioms, theorems, and proofs were a Greek innovation, and the Greeks were primarily geometers. There is no *logical* reason why logic is any more important in geometry than in algebra or arithmetic. Yet it seems that without the explicit statement of assumptions, the parallel postulate of Euclid (discussed in Chapter 10) would never have been questioned. Suppose things had happened that way. Does it follow that non-Euclidean geometry would never have been discovered? How important is non-Euclidean geometry, anyway? What other kinds of geometry do you know about? Is it necessary to be guided by axioms and postulates in order to discover or fully understand, say, the non-Euclidean geometry of a curved surface in Euclidean space? If it is not necessary, what is the value of an axiomatic development of such a geometry?

[1] Charles Coulston Gillispie, *The Edge of Objectivity*, Princeton University Press, 1973.

Answer. Here again we cut to the heart of the usefulness of mathematical abstraction. The primary advantage of mathematics is its use in providing simple models for thought that can be approximated and implemented as what we call "applied mathematics." Now the non-Euclidean geometry that has proved of most use in physics (and it would be hard to name any other area where non-Euclidean geometry has been of any use at all) is part of *differential geometry*, which was created simultaneously with non-Euclidean geometry (and parts of it, even earlier). Its roots are as much algebraic and analytic as geometric. The role of the purer non-Euclidean geometry created by Taurinius, Schweickart, Lobachevskii, and Bólyai, has been to provide a general guide to geometric thought.

1.17. Perminov (*1997*, p. 183) presents the following example of tacit mathematical reasoning from an early cuneiform tablet. Given a right triangle ACB divided into a smaller triangle DEB and a trapezoid $ACED$ by the line DE parallel to the leg AC, such that EC has length 20, EB has length 30, and the trapezoid $ACED$ has area 320, what are the lengths AC and DE? (See Fig. 3.) The author of the tablet very confidently computes these lengths by the following sequence of operations: (1) $320 \div 20 = 16$; (2) $30 \cdot 2 = 60$; (3) $60 + 20 = 80$; (4) $320 \div 80 = 4$; (5) $16 + 4 = 20 = AC$; (6) $16 - 4 = 12 = DE$. As Perminov points out, to present this computation with any confidence, you would have to know exactly what you are doing. What *was* this anonymous author doing?

To find out, fill in the reasoning in the following sketch. The author's first computation shows that a rectangle of height 20 and base 16 would have exactly the same area as the trapezoid. Hence if we draw the vertical line FH through the midpoint G of AD, and complete the resulting rectangles as in Fig. 3, rectangle $FCEI$ will have area 320. Since $AF = MI = FJ = DI$, it now suffices to find this common length, which we will call x; for $AC = CF + FA = 16 + x$ and $DE = EI - DI = 16 - x$. By the principle demonstrated in Fig. 2, $JCED$ has the same area as $DKLM$, so that $DKLM + FJDI = DKLM + 20x$. Explain why $DKLM = 30 \cdot 2 \cdot x$, and hence why $320 = (30 \cdot 2 + 20) \cdot x$.

Could this procedure have been obtained experimentally?

Answer. The first step is obviously to find the average width of the trapezoid, which is the average of DE and AC. As discussed in Chapter 6, the writers of the cuneiform tablets frequently computed the average and half the difference of any two numbers they dealt with. From these two numbers the original numbers could be recovered as the sum and difference. Now the remaining problem is to get half of the difference, which we call the *semi-difference* for convenience. Observe that the area of the rectangle $DKLM$ is 30 times the difference between AC and DE, in other words, 60 times the semi-difference. Hence to get the area of $JCED$, which is the same as the area of $DKLM$, we need to multiply the semi-difference by 60. But, if we add 20 times the semi-difference to the area of $JCED$, we will get 320. That is, 60 + 20 (= 80) times the semi-difference is 320, and therefore the semi-difference is 4. As mentioned, this was a standard technique for finding two numbers $AC = \text{average} + \text{semi} - \text{difference} = 16 + 4 = 20$, and $DE = \text{average} - \text{semi} - \text{difference} = 16 - 4 = 12$.

Whether this answer was obtained experimentally depends on your notion of "experimental." The facts involved are rock-hard geometric truths. Somehow, sometime, someone had to *notice* them, and that noticing was probably an empirical fact at the time, but it was one that accorded with intuition, and it could not have been convincing without some recognition that the two smaller rectangles inside the larger one were obtained by subtracting equal triangles from equal triangles.

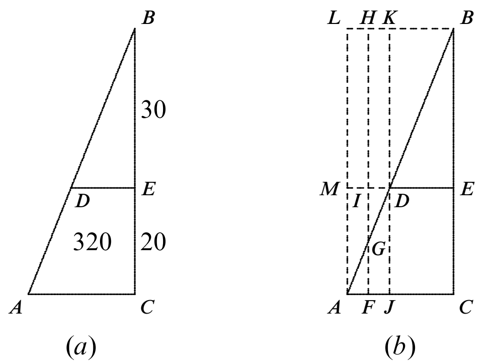


FIGURE 3. (*a*) Line DE divides triangle ABC into triangle DEB and trapezoid $ACED$. (*b*) Line $FGIH$ bisects line AD. Rectangle $FCEI$ has the same area as trapezoid $ACED$, and rectangle $JCED$ equals rectangle $MDKL$.

1.18. A famous example of mathematical blunders committed by mathematicians (not statisticians, however) occurred some two decades ago. At the time, a very popular television show in the United States was called *Let's Make a Deal*. On that show, the contestant was often offered the chance to keep his or her current winnings, or to trade them for a chance at some other unknown prize. In the case in question the contestant had chosen one of three boxes, knowing that only one of them contained a prize of any value, but not knowing the contents of any of them. For ease of exposition, let us call the boxes A, B, and C, and assume that the contestant chose box A.

The emcee of the program was about to offer the contestant a chance to trade for another prize, but in order to make the program more interesting, he had box B opened, in order to show that it was empty. Keep in mind that the emcee *knew* where the prize was and would not have opened box B if the prize had been there. Just as the emcee was about to offer a new deal, the contestant asked to exchange the chosen box (A) for the unopened box (C) on stage. The problem posed to the reader is: Was this a good strategy? To decide, analyze 300 hypothetical games, in which the prize is in box A in 100 cases, in box B in 100 cases (in these cases, of course, the emcee will open box C to show that it is empty), and in box C in the other 100 cases. First assume that in all 300 games the contestant retains box A. Then assume that in all 300 games the contestant exchanges box A for the unopened box on stage. By which strategy does the contestant win more games?

Answer. Since the two strategies are complementary (each wins precisely when the other one loses), the contestant who always retains Box A will win precisely 100 of the 300 games. The contestant who always switches will win the other 200 games. It was smart to switch.

1.19. Explain why the following analysis of the game described in Problem 1.18 leads to an erroneous result. Consider all the situations in which the contestant has chosen box A and the emcee has shown box B to be empty. Imagine 100 games in which the prize is in box A and 100 games in which it is in box C. Suppose the contestant retains box A in all 200 games; then 100 will be won and 100 lost. Likewise, if the contestant switches to box

C in all 200 games, then 100 will be won and 100 lost. Hence there is no advantage to switching boxes.

Answer. This model attempts to define the problem away, It begs the question by assuming from the outset that the prize is equally likely to be on the stage or in the contestant's hand. But it isn't. It is twice as likely that the prize is on stage.

1.20. The fallacy discussed in Problem 1.19 is not in the mathematics, but rather in its application to the real world. The question involves what is known as *conditional probability*. Mathematically, the probability of event E, *given that event F has occurred*, is defined as the probability that E and F both occur, divided by the probability of F. The many mathematicians who analyzed the game erroneously proceeded by taking E as the event "The prize is in box C" and F as the event "Box B is empty." Given that box B has a 2/3 probability of being empty and the event "E and F" is the same as event E, which has a probability of 1/3, one can then compute that the probability of E given F is $(1/3)/(2/3) = 1/2$. Hence the contestant seems to have a 50% probability of winning as soon as the emcee opens Box B, revealing it to be empty.

Surely this conclusion cannot be correct, since the contestant's probability of having chosen the box with the prize is only 1/3 and the emcee can always open an empty box on stage. Replace event F with the more precise event "The emcee has *shown* that box B is empty" and redo the computation. Notice that the emcee is *going* to show that either box B or box C is empty, and that the two outcomes are equally likely. Hence the probability of this new event F is 1/2. Thus, even though the mathematics of conditional probability is quite simple, it can be a subtle problem to describe just what event has occurred. Conclusion: To reason correctly in cases of conditional probability, *one must be very clear in describing the event that has occurred.*

Answer. The problem doesn't leave much for the reader to do. The probability of the new event F is $1/2$, since the emcee is equally likely to show that box B is empty or not (that is, to show that box C is empty). Hence the probability of E given F is $(1/3)/(1/2) = 2/3$.

1.21. Reinforcing the conclusion of Problem 1.20, exhibit the fallacy in the following "proof" that *lotteries are all dishonest*.

Proof. The probability of winning a lottery is less than one chance in 1,000,000 ($= 10^{-6}$). Since all lottery drawings are independent of one another, the probability of winning a lottery five times is less than $\left(10^{-6}\right)^5 = 10^{-30}$. But this probability is far smaller than the probability of any conceivable event. Any scientist would disbelieve a report that such an event had actually been observed to happen. Since the lottery has been won five times in the past year, it must be that winning it is not a random event; that is, the lottery is fixed.

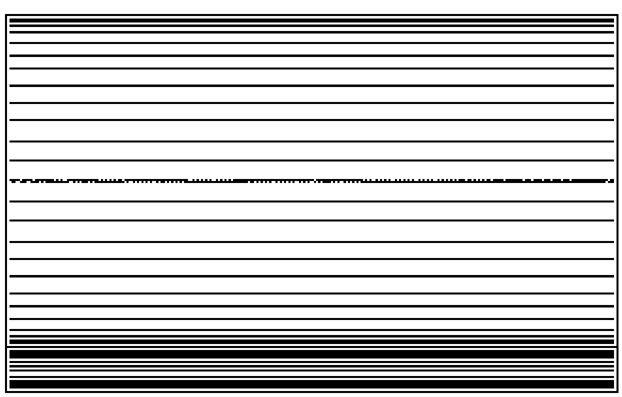
What is the event that has to occur here? Is it "Person A (specified in advance) wins the lottery," or is it "At least one person in this population (of 30 million people) wins the lottery"? What is the difference between those two probabilities? (The same fallacy occurs in the probabilistic arguments purporting to prove that evolution cannot occur, based on the rarity of mutations.)

Answer. The small probability given is the probability that a *specified* ticket will be the winning number. The probability that *a winning ticket will be sold* is much higher, since the number of sets of numbers on the tickets sold is usually a considerable portion (of the order of 10% or even 20%) of the total number.

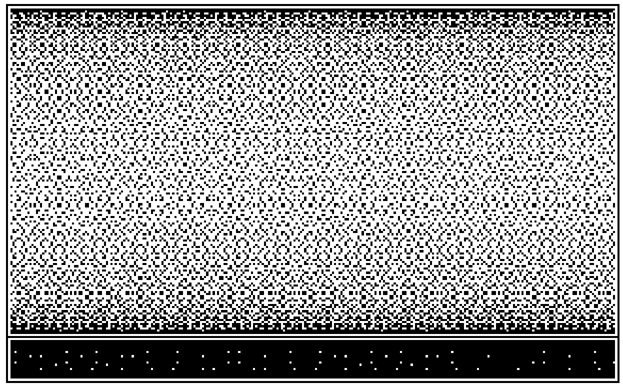
The reference to anti-evolution arguments is included only because they have become very popular of late and embody a very clear and definite fallacy. The probability of a specified organism having a specified mutation is extremely small. The probability that a specified organism has *some* mutation is much higher, since there are so many loci in genetic material where a mutation can occur. Likewise the probability that a given locus will have a mutation within a breeding population is much higher still, especially if the population consists of billions of individual organisms (like bacteria). If a mutation is favorable and confers a reproductive advantage, there will soon be a whole new population of billions of the mutated organism. In geological terms, the time period required is of the order of a few dozen generations, not long at all, even for large animals. A chain of favorable mutations, seen in terms of populations rather than individuals, is not at all an unlikely thing.

1.22. The relation between mathematical creativity and musical creativity, and the mathematical aspects of music itself are a fascinating and well-studied topic. Consider just the following problem, based on the standard tuning of a piano keyboard. According to that tuning, the frequency of the major fifth in each scale should be 3/2 of the frequency of the base tone, while the frequency of the octave should be twice the base frequency. Since there are 12 half-tones in each octave, starting at the lowest A on the piano and ascending in steps of a major fifth, twelve steps will bring you to the highest A on the piano. If all these fifths are tuned properly, that highest A should have a frequency of $\left(\frac{3}{2}\right)^{12}$ times the frequency of the lowest A. On the other hand, that highest A is seven octaves above the lowest, so that, if all the octaves are tuned properly, the frequency should be 2^7 times as high. The difference between these two frequency ratios, $7153/4096 \approx 1.746$ is called the *Pythagorean comma*. (The Greek word *komma* means a break or cutoff.) What is the significance of this discrepancy for music? Could you hear the difference between a piano tuned so that all these fifths are exactly right and a piano tuned so that all the octaves are exactly right? (The ratio of the discrepancy between the two ratios to either ratio is about 0.01%.)

Answer. The last sentence gives the answer: One can feel a difference in tonal quality between the two, but it is difficult to call it a difference in *pitch*. However, when the two tones are played together, the result is a nerve-shattering twitter. Since it is difficult to impart sound in a book, a visual representation will have to suffice. When *Mathematica* is used to produce 2-second long sinusoidal waves having these frequencies, via the command `Play[{Sin[2π a t], Cos[2π b t]}, {t,0,2}]`, where $a = 3520$ for the high A when the piano is tuned in octaves and $b = 29229255/8192 \approx 3568.02$ for the same high A when it is tuned in fifths, the resulting sound waves are depicted graphically as follows:

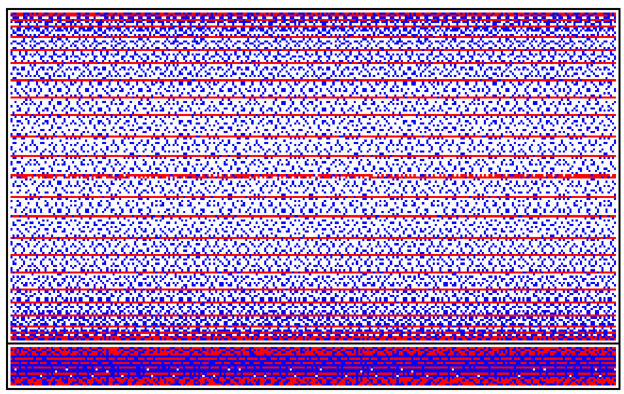


Octaves ($a = 3520$)



Fifths ($b = 3568.02$)

When the two tones are played together, the picture is as follows. (You can see the conflict here.)



The tones superimposed

1.23. What meaning can you make of the statement attributed to the French poet Sully (René François Armand) Prudhomme (1839–1907), "Music is the pleasure the soul experiences from counting without realizing it is counting"?

Answer. The unconscious counting referred to is probably the rhythmic alternation of stressed and unstressed beats and the syncopation effects that result when two different patterns of rhythm are superimposed, such as (my personal favorite) 3 beats against 4. This kind of counting can be felt directly by those who "have rhythm" (most people), without the need for conscious use of numbers.[2] In fact, it is extremely difficult to *count* three beats against four, but I have found that a whole classroom of students can learn to beat three against four with their two hands in less than a minute by simply listening and imitating, getting a feel for the rhythm.

There is a second kind of periodicity that Prudhomme probably *didn't* have in mind, the kind that is sensed as *pitch* and results in the tones that result from string and wind instruments in the form of chords and counterpoint. These frequencies are too rapid to be

[2] Music teachers have traditionally told their pupils to count. Fortunately those among them who were fated to become real musicians managed to ignore this advice.

counted, but the mathematical analysis of them, starting with the problem of the vibrating string in the eighteenth century, reveals an intricate mathematical structure in the theory of sound in general.

CHAPTER 2

Mathematical Cultures I

2.1. Does mathematics realize Plato's program of understanding the world by contemplating eternal, unchanging forms that are perceived only by reason, not by the senses?

Answer. It would be presumptuous to claim anything like a definitive answer to this question. Arithmetic and geometry deal with *ideas* that correspond to *relations* in the physical world. Are *relations* real, physical things, like the elementary particles that we imagine the world is made of? If not, how does it happen that these ideas in human minds correspond to the physical world so well that they enable us to predict and sometimes control what happens? If relations are in some sense a real part of the world, what does their reality consist of? I am very far from regarding myself as a Platonist—I think his metaphysics is a hopeless muddle—and yet there is a certain plausibility about Plato's idea of linking the observed world with a world that exists only in thought. A case in point is the notion of energy, which is a mathematical function of position and velocity in classical mechanics. Is energy real? Some would say it is the *only* reality, that everything is a form of energy. Yet one does not observe energy directly: To find its value for a given physical system, you observe the position and velocity of a mass, or measure the strength of electric and magnetic fields, or take the temperature of a sample of matter. What is the underlying reality? The energy or these observables? Energy is of supreme importance in physics: Where would physics be without the law of conservation of energy, or the matter-energy equivalence expressed by $E = mc^2$? With these ruminations, I leave the question to the student.

2.2. To what extent do the points of view expressed by Hamming and Hardy on the value of pure mathematics reflect the nationalities of their authors and the prevailing attitudes in their cultures? Consider that unlike the public radio and television networks in the United States, the CBC in Canada and the BBC in Britain do not spend four weeks a year pleading with their audience to send voluntary donations to keep them on the air. The BBC is publicly funded out of revenues collected by requiring everyone who owns a television set to pay a yearly license fee.

Answer. Public opinion, at least in my own informal, unsystematic sample of US citizens, divides along conservative-liberal lines (as those words are currently understood), with the liberals supporting an "elitist" view that the government has an obligation to raise the general level of culture of its citizens and conservatives supporting the "populist" view that people ought not to be taxed to support arts, entertainment, and research that does not bring them any direct benefits. I leave to any European and Canadians I may be fortunate enough to number among my readers to give their view of the contrast between their own countries' systems and the system in the United States, and I would not presume to answer for them. I personally like what is on public television and radio and have almost no use for the AM dial on my radio, but I do appreciate the libertarian objection to publicly funded broadcasts. If the technology existed for sending out broadcast signals in code and allowing each broadcaster to charge a monthly fee for decoding technology, I would be delighted. Then each of us could pay the full price for the news and entertainment of his or her own choice.

2.3. In an article in the *Review of Modern Physics*, **51**, No. 3 (July 1979), the physicist Norman David Mermin (b. 1935) wrote, "Bridges would not be safer if only people who knew the proper definition of a real number were allowed to design them" (quoted by Mackay, *1991*, p. 172). Granting that at the final point of contact between theory and the physical world, when a human design is to be executed in concrete and steel, every number is only an approximation, is there any value for science and engineering in the concept of an infinitely precise real number? Or is this concept only for idealistic, pure mathematicians? (The problems below may influence your answer.)

Answer. The question is phrased in such a way that a fuller answer can be given after the following questions are answered. In the meantime, see the comments on the usefulness of $\sqrt{2}$ in Problem 1.10. May we presume that the Tacoma Narrows Bridge was designed by a person who *did* know the proper definition of a real number?

2.4. In 1837 and 1839 the crystallographer Auguste Bravais (1811–1863) and his brother Louis (1801–1843) published articles on the growth of plants.[1] In these articles they studied the spiral patterns in which new branches grow out of the limbs of certain trees and classified plants into several categories according to this pattern. For one of these categories they gave the amount of rotation around the limb between successive branches as $137^\circ\, 30'\, 28''$. Now, one could hardly measure the limb of a tree so precisely. To measure within 10° would require extraordinary precision. To refine such crude measurements by averaging to the claimed precision of $1''$, that is, $1/3600$ of a degree, would require thousands of individual measurements. In fact, the measurements were carried out in a more indirect way, by counting the total number of branches after each full turn of the spiral. Many observations convinced the brothers Bravais that normally there were slightly more than three branches in two turns, slightly less than five in three turns, slightly more than eight in five turns, and slightly less than thirteen in eight turns. For that reason they took the actual amount of revolution between successive branches to be the number we call $1/\Phi = (\sqrt{5} - 1)/2 = \Phi - 1$ of a complete (360°) revolution, since

$$\frac{3}{2} < \frac{8}{5} < \Phi < \frac{13}{8} < \frac{5}{3}.$$

Observe that $360^\circ \div \Phi \approx 222.4922359^\circ \approx 222^\circ\, 29'\, 32'' = 360^\circ - \left(137^\circ\, 30'\, 28''\right)$. Was there scientific value in making use of this *real* (infinitely precise) number Φ even though no actual plant grows exactly according to this rule?

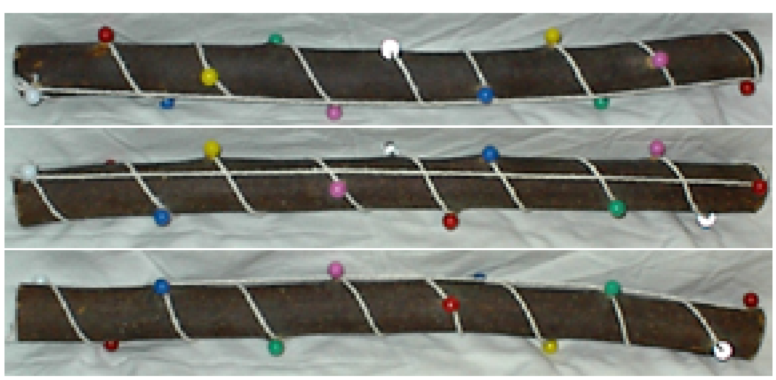
Answer. Obviously, I think the answer is "Yes." The positive value that I see is in obtaining a single number for the amount of rotation between branches, to which the actual numbers observed in many physical specimens can be regarded as approximating. As in Problem 1.10, the number Φ provides a single, simple mathematical model that fits approximately a large number of observable cases and makes some sense out of what would otherwise be a chaotic jumble of observations and numbers. The number Φ would never

[1] See the article by I. Adler, D. Barabe, and R. V. Jean, "A history of the study of phyllotaxis," *Annals of Botany*, **80** (1997), 231–244, especially p. 234. The articles by Auguste and Louis Bravais are "Essai sur la disposition générale des feuilles curvisériées," *Annales des sciences naturelles*, **7** (1837), 42–110, and "Essai sur la disposition générale des feuilles rectisériées," *Congrès scientifique de France*, **6** (1839), 278–330.

have been discovered by measurement, but in this case the counting technique that the Bravais brothers used (like the technique for discovering the ratio between a month and a year) showed near matches between revolutions and numbers of branches at successive numbers in the "Fibonacci" sequence 1, 2, 3, 5, 8, 13,..., where each term is the sum of its two predecessors. It is well-known that the limit of the ratio of each term to its predecessor is Φ. Moreover, as the Bravais brothers found, if you start a "Fibonacci" sequence with the same rule but different initial numbers, for example, 1, 3, 4, 7, 11, 18,..., *there are actual families of plants that grow according to the corresponding pattern* . Thus the mathematics has more than a superficial connection with the phenomenon being studied.

2.5. Plate 4 shows a branch of a flowering crab apple tree from the author's garden with the twigs cut off and the points from which they grew marked by pushpins. The "zeroth" pin at the left is white. After that, the sequence of colors is red, blue, yellow, green, pink, clear, so that the red pins correspond to 1, 7, and 13, the blue to 2 and 8, the yellow to 3 and 9, the green to 4 and 10, the pink to 5 and 11, and the clear to 6 and 12. Observe that when these pins are joined by string, the string follows a helical path of nearly constant slope along the branch. Which pins fall nearest to the intersection of this helical path with the meridian line marked along the length of the branch? How many turns of the spiral correspond to these numbers of twigs? On that basis, what is a good approximation to the number of twigs per turn? Between which pin numbers do the intersections between the spiral and the meridian line fall? For example, the fourth intersection is between pins 6 and 7, indicating that the average number of pins per turn up to that point is between $\frac{6}{4} = 1.5$ and $\frac{7}{4} = 1.75$. Get upper and lower estimates in this way for all numbers of turns from 1 to 8. What are the narrowest upper and lower bounds you can place on the number of pins per turn in this way?

Answer. This question gives me a chance to insert some clearer pictures of the branch in question, as the one in Plate 4 was rather blurry. Here are three views of the branch, rotated through a right angle from one to the next.



If you trace around the spiral, counting both the number of branches and the number of turns as you go, you can log your "trip" from the white pin at the start to the red pin at the end in the following table:

Turns	**Branches**
0	0 (white)
	1 (red)
1	
	2 (blue)
	3 (yellow)
2	
	4 (green)
3	
	5 (pink)
	6 (clear)
4	
	7 (red)
	8 (blue)
5	
	9 (yellow)
6	
	10 (green)
	11 (pink)
7	
	12 (clear)
8	13 (red)

From this table, by comparing the number of branches indicated just before and just after each complete revolution, you can get the following seven inequalities for the number r, the average number of branches per turn:

$$\begin{aligned}
1 <{} & r < 2\\
1.5 = \frac{3}{2} <{} & r < 2\\
1.33 \approx \frac{4}{3} <{} & r < \frac{5}{3} \approx 1.67\\
1.5 = \frac{6}{4} <{} & r < \frac{7}{4} = 1.75\\
1.6 = \frac{8}{5} <{} & r < \frac{9}{5} = 1.8\\
1.5 = \frac{9}{6} <{} & r < \frac{10}{6} \approx 1.67\\
1.57 \approx \frac{11}{7} <{} & r < \frac{12}{7} \approx 1.71
\end{aligned}$$

From these inequalities, choosing the strongest on both sides, we find $1.6 < r < 1.67$. The last line of the table shows that $r \approx \frac{13}{8} = 1.625$.

2.6. Suppose that the pins in Plate 4 had been joined by a curve winding in the opposite direction. How would the numbers of turns of the spiral and the number of pins joined compare? What change would occur in the slope of the spiral?

Answer. If you can visualize this, you have very good geometric intuition. The table given above gets replaced by the following table.

Turns	**Branches**
0	0 (white)
	1 (red)
	2 (blue)
1	
	3 (yellow)
	4 (green)
	5 (pink)
2	
	6 (clear)
	7 (red)
3	
	8 (blue)
	9 (yellow)
	10 (green)
4	
	11 (pink)
	12 (clear)
5	13 (red)

As a result, one gets the following set of inequalities for the number of branches per turn when we spiral in this direction:

$$\begin{aligned} 2 <{} & r < 3 \\ 2.5 = \frac{5}{2} <{} & r < \frac{6}{2} = 3 \\ 2.33 \approx \frac{7}{3} <{} & r < \frac{8}{3} \approx 2.67 \\ 2.5 = \frac{10}{4} <{} & r < \frac{11}{4} = 2.75 \end{aligned}$$

That is to say, each of our estimates is "moved up one," but the series is shortened, since only five turns of the spiral occur. The final result shows that we get about $\frac{13}{5} = 2.6$ branches per turn this way. In terms of theory, we ought to expect this. If indeed the angle of turning between successive branches is $1/\Phi$ of a revolution, then when we go the other way, the angle ought to be $1-1/\Phi$ of a revolution between successive branches, and so the number of branches per turn should be the reciprocal of this number, that is, $\Phi/(\Phi-1)$. But since Φ satisfies the equation $\Phi - 1 = 1/\Phi$, this number should be Φ^2, which by this equation is $\Phi + 1$. As a corollary, the slope of the spiral increases by a factor of Φ (approximately $\frac{8}{5}$ as far as our sample shows).

2.7. With which of the two groups of people mentioned by Plato do you find yourself more in sympathy: the "practical" people, who object to being taxed to support abstract speculation, or the "idealists," who regard abstract speculation as having value to society?

Answer. Again, this is purely a matter of political and social preference. I personally believe strongly, as Plato did, in the value of pure thought and all the other accoutrements of what used to be called a liberal education. On the other hand, I am reluctant to trick people into supporting this cause by urging spurious arguments that engineers are better engineers if they know history. It is very difficult to measure what economists call the marginal value of time spent in history courses in place of time spent in engineering courses and next to impossible to prove that that marginal value is positive rather than negative. I prefer to use persuasion to convince people that *they will feel rewarded* by learning abstract mathematics, philosophy, history, music, poetry and so on. They *may* make some professional use of these subjects, but that is not the aim of teaching them.

2.8. The division between the practical and the ideal in mathematics finds an interesting reflection in the interpretation of what is meant by solving an equation. Everybody agrees that the problem is to find a number satisfying the equation, but interpretations of "finding a number" differ. Inspired by Greek geometric methods, the Muslim and European algebraists looked for algorithms to invert the operations that defined the polynomial whose roots were to be found. Their object was to generate a sequence of arithmetic operations and root extractions that could be applied to the coefficients in order to exhibit the roots. The Chinese, in contrast, looked for numerical processes to approximate the roots with arbitrary accuracy. What advantages and disadvantages do you see in each of these approaches? What would be a good synthesis of the two methods?

Answer. There is beauty in both the abstract approach and in the numerical approach. The abstract approach (Galois theory) can play a role here in relation to the numerical approach analogous to the role I have envisioned above for mathematics in general in relation to science. That is, it can suggest what is possible and act as a guide for numerical work.

2.9. When a mathematical document such as an early treatise or cuneiform tablet contains problems whose answers "come out even," should one suspect or conclude that it was a teaching device—either a set of problems with simplified data to build students' confidence or a manual for teachers showing how to construct such problems?

Answer. These documents are seldom "applied mathematics." It is overwhelmingly likely that they were used for teaching and conveying the techniques of mathematical thought to pupils who were to become "mathematicians" (scribes and overseers). That is why the problems on them so often come out even. When we give examples, we like to keep them as simple as possible.

2.10. From what is known of the Maya codices, is it likely that they were textbooks intended for teaching purposes, like many of the cuneiform tablets and the early treatises from India, China, and Egypt?

Answer. The Maya codices do not seem to contain any instructions about methods of computation. Instead, they record observations and recount the Maya mythology relating to the heavenly bodies. Of course, one should not forget the burning of the Maya books. We are working from a very small sample here.

2.11. Why was the Chinese encounter with the Jesuits so different from the Maya encounter with the Franciscans? What differences were there in the two situations, and what conditions account for these differences? Was it merely a matter of the degree of zeal that inspired Diego de Landa and Matteo Ricci, or were there institutional or national differences between the two as well? How much difference did the relative strength of the Chinese and the Maya make?

Answer. Archaeological evidence indicates that the Maya civilization was in decline, probably because of climatic disasters (droughts and famines) many centuries before the Spanish arrived. Moreover, their population and resources were much smaller. It is difficult to believe that the *conquistadores* could have done to the Chinese, with their long continuous tradition of a literate civil service, what they succeeded in doing to the Maya. Of course, history bears out this conclusion. Almost the last vestige of European dominance in Asia disappeared when the British returned Hong Kong to Chinese rule in 1999. (The precarious independence of Taiwan, defended by the threat of American military intervention, is the very last.) In contrast, the European character of the Americas is now so deeply rooted that it is impossible to imagine its disappearing any time soon.

CHAPTER 3

Mathematical Cultures II

3.1. Compare the way in which mathematicians have been supported in various societies discussed in this chapter. If you were in charge of distributing the federal budget, how high a priority would you give to various forms of pure and applied research in mathematics? What justification would you give for your decision? Would it involve a practical "payoff" in economic terms, or do you believe that the government has a responsibility to support the creation of new mathematics, without regard to its economic value?

Answer. It is revealing that Plato thought it the duty of the state (the *polis* or city-state of his time) to support research into geometry. Plato was a member of the aristocracy of his time and took it for granted that his class had the right to rule and to spend public funds in accordance with its own priorities. As a citizen of a modern democracy, I do not believe in any such absolute right, only the right of the elected government to govern. But that comes down to the same thing in the end. Like Plato, I consider the elements of a liberal education to be of intrinsic value. However, I much prefer that they be funded privately, through tuition charged to students who benefit from the education. I also believe that most of the research that plays such a large role in promotion, tenure, and salary decisions at our universities is a distraction from the teaching duties of professors. Except for those who can obtain government or private funding for their research, professors should do a *modest* amount of research, and it should be directly related to their teaching.

Of more interest are the early European universities, where people went to become experts in the classical quadrivium and trivium and such areas as theology. There was a rather narrow orthodoxy in this last subject, and it developed into an elaborate system of doctrine irrelevant to the lives of most people. Exactly the same is true of mathematics today, and there are people much puffed up about their mastery of such arcane areas as topological semi-groups. They expect to be, and often are, respected for having spent large parts of their life learning about these ethereal entities. Looked at from a commonsense point of view, it is rather mysterious that one can "sell" such products to the public, but universities have managed to do so for many centuries. I often wonder if the whole economy of higher education has been based on a public misunderstanding. Did people in the Middle Ages believe that the doctors of divinity bestowed some mysterious benefit on society with their word-spinning disputes? Do people believe that today about, say, deconstructionism? I do not wish this argument to be interpreted as an anti-intellectual blast against abstract knowledge. As I have stressed in the text and in other parts of this answer manual, I consider the ability merely to understand the world to be of great value. My point is that A's understanding of the world through a liberal education is not of direct benefit to B. If I ask B to support higher education, I want it to be on the basis that B gets a higher education that really does enable B to understand the actual world of his or her experience better, not the dream world of some of the more bizarre academic fads.

3.2. Why is Seki Kōwa the central figure in Japanese mathematics? Are comparisons between him and his contemporary Isaac Newton justified?

Answer. Like Archimedes and Newton, Seki Kōwa mastered a wide range of mathematics and introduced new methods that changed the way his successors thought about mathematical problems. The scale of his achievements is not really comparable with Newton, however, since the mathematics that he inherited from his predecessors was far less elaborate.

3.3. What is the justification for the statement by the historian of mathematics T. Murata that Japanese mathematics was not a science but an art?

Answer. It is very tempting to compare Japanese mathematics of the eighteenth century with the delicate Japanese paintings of the same period. Indeed, the two almost merge in the *Sangaku* shown in Plate 2. The criteria for creativity in both cases seem to have been esthetic. Little if any of this mathematics was aimed at the needs of physical science or commerce or technology.

3.4. Why might Seki Kōwa and other Japanese mathematicians have wanted to keep their methods secret, and why did their students, such as Takebe Kenkō, honor this secrecy?

Answer. In the case of the Japanese mathematicians, I suspect a mystical, philosophical, or religious motive, which the students would honor out of respect for the *sensei*. Mathematics seems to have been a leisure-time activity, not a profession, in Japan. In contrast, mathematics among the sixteenth-century Italians had both intellectual and economic aspects.

3.5. For what purpose was algebra developed in Japan? Was it needed for science and/or government, or was it an "impractical" liberal-arts subject?

Answer. Almost certainly, it was "liberal-artsy." At least, none of the historians I have read on this subject have pointed out any economic uses of the mathematics.

3.6. Dante's final stanza, quoted in the text, uses the problem of squaring the circle to express the sense of an intellect overwhelmed, which was inspired by his vision of heaven. What resolution does he find for the inability of his mind to grasp the vision rationally? Would such an attitude, if widely shared, affect mathematical and scientific activity in a society?

Answer. Dante found his peace in surrender to the love of God. This outcome may be morally higher than continued efforts to master the subject intellectually, but it amounts to a decision to stop thinking about the problem, not a solution of the problem. One may certainly debate how long one should continue to think about an unresolved problem. At some point, it certainly must become an unprofitable use of one's time. Whether religious meditation is more profitable is a question that I leave to the reader.

3.7. One frequently repeated story about Christopher Columbus is that he proved to a doubting public that the Earth was round. What grounds are there for believing that "the public" doubted this fact? Which people in the Middle Ages would have been likely to believe in a flat Earth? Consider also the frequently repeated story that people used to believe the stars were near the Earth. How is that story to be reconciled with Ptolemy's assertion that it was acceptable to regard Earth as having the dimensions of a point relative to the stars?

Answer. The answer must be understood within the context of education in Columbus' day. The overwhelming majority of people were illiterate and ignorant in a way that we can hardly imagine today. What they believed about the shape of the earth is not easy to ascertain, nor is it a question of any importance. But educated people during the Middle Ages, who read Ptolemy, certainly knew that the universe was large—not as large as we now know it to be, but very large nevertheless—and that the earth was spherical.

3.8. What are the possible advantages and disadvantages of eliminating or greatly reducing the volume of journals, placing all articles on electronic files that can be downloaded from various information systems?

Answer. Having experienced the joy of getting information from the World Wide Web that would have taken weeks to acquire through Interlibrary Loan, and having used search engines to locate sources whose existence I would never have suspected, I am an enthusiastic proponent of putting absolutely everything on the Web—photostatic copies of ancient manuscripts, texts of existing journals, and all new journals and books. Think of the trees we can save, not to mention the overstuffed, inaccessible warehouses where most libraries are now forced to put all their antiquarian materials (from those dim ages of the distant past, such as the 1970s). But the primary benefit is the rapid, universal accessibility of information stored in this way. The only disadvantage is in the progress of technology, which will force the continuous acquisition of new methods of storage and retrieval. However, conversions to updated versions of the storage and retrieval software will no doubt be accomplished by computers and should not be a serious problem.

3.9. Mathematical research is like any other commercial commodity in the sense that people have to be paid to do it. We have mentioned the debate over taxing the entire public to support such research and asked the student to consider whether there is a national interest that justifies this taxation. A similar taxation takes place in the form of tuition payments to American universities. Some of the money is spent to provide the salaries of professors who are required to do research. Is there an educational interest in such research that justifies its increased cost to the student?

Answer. I believe the answer to this question is contained in my answers to several other questions above. It is a qualified, "Yes." Professors should do research *related to their teaching*, and that research should be funded out of tuition. Research into esoterica of interest only to the initiates in a given area should not be funded out of money students have paid in order to be taught.

CHAPTER 4

Women Mathematicians

4.1. In the late fourth and early fifth centuries the city of Alexandria, where Hypatia lived, was divided into Christian, Jewish, and pagan cultures. Is it merely a random event that the only woman mathematician of the time in this city with a long history of scholarship happened to come from the pagan culture?

Answer. In the modern world both Jewish and Christian cultures have liberated women and encouraged them to develop their talents. In the early Christian world, women played a large and vital role in the community and were often prominent leaders. However, the community itself did not produce any scientists, male or female; scientific curiosity was not among its virtues. At the time when Hypatia lived, Christianity was developing a patriarchal system of governance that relegated women to an auxiliary role for many centuries. The pagan culture was apparently more egalitarian in this regard, but it was rapidly losing out to Christianity, as shown by the termination of the Olympic Games by the Emperor Theodosius in the year 395.

4.2. Compare the careers of Charlotte Angas Scott and Sof'ya Kovalevskaya. In what aspects were they similar? What significant differences were there? Were these differences due to the continental circles in which Kovalevskaya moved compared to the Anglo-American milieu of Scott's career? Or were they due to individual differences between the two women?

Answer. The most prominent difference that springs to mind is Kovalevskaya's marriage of convenience, the only route she was able to find to pursue the career she wanted. While marriages of convenience were known in Britain and France, I cannot think of one that was entered into merely as a way of pursuing a career. As Elizaveta Fedorovna Litvinova (1845–1919), another Russian woman who went to the West (Zürich) to obtain a degree, explained, "You don't have to be a genius to understand this, but you do have to be a Russian." Of course, there were individual differences between the two women, hard for us to judge, who never knew either woman. Still, from the public record that remains of them, Scott seems to have been a much more even-tempered, phlegmatic type than the volatile Kovalevskaya.

4.3. Choose two women mathematicians, either from among those discussed in this chapter or by going to a suitable website. Read brief biographical sketches of them. Then try to match each woman with a comparable male mathematician from the same era and country. Compare their motives for studying mathematics if any motives are given, the kind of education they received, the journals where they published their work, and the kind of academic positions they occupied.

Answer. This question, now that I have tried to answer it, seems far less useful than it seemed when I posed it. The reason is that individual differences simply swamp all other differences when the sample is only two people. Only in terms of statistics are any generalizations possible. Be that as it may, here are my two examples:

Case 1: Julia Bowman Robinson (1919–1985) and Norman Steenrod (1910–1971). I matched these two only because they were nearly contemporaries, and because both were born in the American Midwest, Robinson in Saint Louis, Missouri, and Steenrod in Dayton, Ohio.

Julia Bowman's mother died when Julia was only two years old, and she and her older sister Constance (the author Constance Bowman Reid, whose biographies of famous mathematicians—Julia Bowman Robinson among them—have won great acclaim) were sent to live in Arizona, where their father joined them a year later, after his remarriage. Julia had serious illnesses as a child and was tutored at home for a while. In high school, she was the only girl in the more advanced science and mathematics courses. She went to San Diego State University. The Great Depression of the 1930s bankrupted her father, and he committed suicide during her second year of college. Supported by her aunt and her sister, she transferred to the University of California at Berkeley for her senior year, and blossomed as a student. There she met and married Raphael Robinson, a professor of mathematics. Temporarily, because of anti-nepotism rules, her marriage derailed her academic career. A miscarriage and her doctor's advice not to have children devastated her, but she turned her attention to mathematics and became one of the quartet of brilliant mathematicians who together solved Hilbert's Tenth Problem, proving that no algorithm could exist to determine whether a Diophantine equation has solutions. (The other three were Hilary Putnam, Martin Davis, and Yuri Matiyasevich. Matiyasevich furnished the final step in 1970.) She became the first woman officer of the American Mathematical Society, the second to give one of its Colloquium letters, and its first president in 1982. In 1976 she became the first woman elected to the Academy of Sciences. Matiyasevich later recalled that she had insisted on having her name written out in full on an expository publication about Hilbert's Tenth Problem, to avoid being confused with John Robinson and George Robinson, names that occurred in close conjunction with hers in a number of contexts; ironically, he noted, he himself never realized that the J. Robinson whom he knew of from a paper on game theory was in fact the Julia Robinson that he knew from the Hilbert Problem. (She had published that paper in 1951 while working for the RAND corporation.) She died (of leukemia) at the age of 65.

Norman Steenrod was the brilliant child of school teachers, who completed high school at the age of 15 and enrolled in Miami of Ohio at the age of 17. He transferred to the University of Michigan, where he made his first acquaintance with topology, the subject that was to be at the heart of his career, graduating in 1932. Unfortunately, he did not get any fellowship to graduate school. He went back home, worked on his own, and wrote a good paper in topology. On that basis, he received an offer from Harvard, where he went for a year before joining his former Michigan professor Wilder at Princeton. There he became a student of Solomon Lefschetz and soon obtained the PhD. He married in 1938, and taught from 1939 through 1942 at the University of Chicago. He then taught for five years at the University of Michigan, before coming to Princeton in 1947. He was famous for his courses in topology (I myself attended one of them in 1963) and for his research and expository work in the topology of fiber bundles. He was invited to give an AMS Colloquium Lecture in 1957 and eventually became a member of the Academy of Sciences.

In many ways these two mathematicians are quite comparable. Both, near the end of their careers, abandoned pure research (Robinson because she was busy as president of the

AMS, Steenrod because he realized that his best ideas were those he had had in the past—at least so I was told by another Princeton faculty member.) One must say, however, that it is very unlikely that Robinson would have been given the offers that Steenrod received. Harvard and Princeton in those days, the best possible places for mathematicians, were not congenial to the idea of recruiting women. Fortunately it is well recognized that Berkeley is quite comparable to Harvard and Princeton, and Berkeley apparently made no issue of Robinson's gender—except for that anti-nepotism rule mentioned above.

Case 2: Olga Aleksandrovna Ladyzhenskaya (1922–2004) and Georgii Dmitrievich Suvorov (1919–1984). These two Russians lived in rather different parts of their vast country. Ladyzhenskaya (pronounced La-DIH-zhenskaya) graduated from Moscow University in 1947 and did graduate work at Leningrad University, where she taught. After 1949 she was also a member of the Academy of Sciences (Steklov) Mathematics Institute, Leningrad Branch. Her work in partial differential equations is widely recognized as among the best work done during the twentieth century in this area.

Suvorov, although born in Saratov, in western Russia, graduated from the University of Tomsk, in Siberia, in 1941. I do not know what if any military service he performed during the titanic struggle against the Nazis, but it is unlikely that he escaped military service entirely. He completed his graduate work at Tomsk in 1949 (after the war) and taught there until 1965, at which point he moved to Donetsk, in the Ukraine, to head a department in the Academy of Sciences Institute of Applied Mathematics and Mechanics. He is famous for his outstanding work in conformal mapping.

There are many similarities between these two mathematicians, in terms of the acclaim their work received from specialists in their areas, although Ladyzhenskaya is certainly better known in general, justifiably so in terms of the quantity of work produced. The Soviet publication *Mathematics in the USSR after Forty Years, 1917–1957* listed 32 papers by Ladyzhenskaya between 1950 and 1957, ten by Suvorov from 1948 to 1957.

4.4. How do you account for the fact that a considerable percentage (compared to their percentage of the general population) of the women studying higher mathematics in the United States during the 1930s were Roman Catholic nuns? (Some of these nuns produced mathematical research of high quality, for example, Sister Mary Celine Fasenmyer (1906–1996).)

Answer. First, let us give an idea of the impact of nuns on graduate work in mathematics. Sister Miriam Cooney, who received the Ph.D. in 1969 from the University of Chicago for a dissertation in algebra, conducted a study of over 100 nuns with PhDs in mathematics. Considering how small a portion of the population nuns amounted to, even when they were relatively numerous, that is an amazingly large number of mathematicians. There are probably many causes for this phenomenon, all connected with the fact that nuns tend to lead a structured life, in which teaching and studying often plays a part, and that their services were much in demand as teachers in Catholic schools and colleges. Of course, there is a long tradition of the outstanding scholarly nun, going back at least as far as Hildegard of Bingen (1098–1179) and Sor Juana Inés de la Cruz (1648–1695), so these women had very good role models.

4.5. What were the advantages and disadvantages of marriage for a woman seeking an academic career before the twentieth century? How much of this depended on the particular choice of a husband at each stage of the career? The cases of Mary Somerville, Sof'ya Kovalevskaya, and Grace Chisholm Young will be illuminating, but it will be useful to seek more detailed sources than the narratives above.

Answer. This question is impossible to answer accurately, since we will never know how many talented women were simply extinguished intellectually by getting married and transferring all their energies to domestic concerns. A comparison of the relative number of single women among scientists and single women in the general population, if one exists, might be of some help in answering this question. As noted in the text, Mary Somerville's first husband was a definite hindrance to her career, but her second husband seems to have helped. Kovalevskaya's husband, who began by attempting to help her, ultimately became a burden to her. Had he lived, I doubt if she would have achieved the fame that came to her as a good mathematicians who was also a respectable widow with a daughter. Grace Chisholm Young, as noted, hid her talents behind her husband, at least at first.

Thus, the disadvantages of marriage can be considerable. On the other hand, most people do want to have children, and a marriage between mathematicians seems to increase greatly the chances that the children will be scholars. The Youngs are one example. Another, from Russia, is the marriage of Pelageya Yakovlevna Polubarinova (1899–1999) and Nikolai Evgrafovich Kochin (1901–1944), whose daughter Nina Nikolaevna Kochina wrote an article (`http://www-sbras.nsc.ru/HBC/2000/n09/f9.html`) examining the fact that, despite what appeared to be full access to educational opportunities in the Soviet Union, the number of women in both government and science there was much smaller than the number of men. (The article is in Russian, and "HBC" is really "NVS," standing for the Russian phrase "science in Siberia.")

The best "more detailed source" is *Women of Mathematics. A Biobibliographic Sourcebook*, edited by Louise S. Grinstein and Paul J. Campbell. Greenwood Press, Westport, CT, 1987.

4.6. How big a part did chance play in the careers of the early women mathematicians? (The word *chance* is used advisedly, rather than *luck*, since the opportunities that came for Sof'ya Kovalevskaya and Anna Johnson Pell Wheeler were the result of tragic misfortunes to their husbands.)

Answer. Since there were no regular channels for women to become mathematicians until a century ago, all were forced to make their own routes, taking advantage of whatever opportunities came their way. However, those opportunities were often mere chance remarks and encouragement from mentors or friends and family, which turned a young woman's mind to the possibility of becoming a mathematician.

4.7. How important is (or was) encouragement from family and friends in the decision to study science? How important is it to have a mentor, an established professional in the same field, to help orient early career decisions? How important is it for a young woman to have an older woman as a role model? Try to answer these questions along a scale from "not at all important" through "somewhat important" and "very important" to "essential." Use the examples of the women whose careers are sketched above to support your rankings.

Answer. If being a mathematician is ever going to be a normal thing for a woman, not at all marking her as different, such mentoring needs to be there. It is almost as important for these role models to be apparent to the public in general as for them to be known to the aspiring young mathematician. When a woman can tell people of her plan to become a mathematician without getting a curious look that says, "How unusual!" we will know we have nearly succeeded in "leveling the playing field." I think we are nearly there; general sensitivity to gender stereotypes has increased greatly of late. In the meantime, I would regard answer all of these questions as "very important." Of course a woman can get along without them if necessary, but there is no doubt that they make things easier.

4.8. Why were most of the women who received the first doctoral degrees in mathematics at German universities foreigners? Why were there no Germans among them? In his lectures on the development of nineteenth-century mathematics (*1926*, Vol. 1, p. 284), Klein mentions that a 17-year-old woman named Dorothea Schlözer (apparently German, to judge by the name) had received a doctorate in economics at Göttingen a full century earlier.

Answer. Probably the authorities in charge of such things at the universities and in the government thought they could control the numbers of foreign women entering German universities and wouldn't have to let them into the workplace in Germany.

4.9. How strong are the "facts" that Loria adduces in his argument against admitting women to universities? Were all the women discussed here encouraged by their families when they were young? Is it really true that it is impossible to "fix with precision" the original contributions of Sophie Germain and Sof'ya Kovalevskaya? You may wish to consult biographies of these women in which their correspondence is discussed. Would collaboration with other mathematicians make it impossible to "fix with precision" the work of any male mathematicians? Consider also the case of Charlotte Angas Scott and others. Is it true that they were exhausted after finishing their education?

Next, consider what we may call the "honor student" fallacy. Universities select the top students in high school classes for admission, so that a student who excelled the other students in high school might be able at best to equal the other students at a university. Further selections for graduate school, then for hiring at universities of various levels of prestige, then for academic honors, provide layer after layer of filtering. Except for an extremely tiny elite, those who were at the top at one stage find themselves in the middle at the next and eventually reach (what is ideally) a level commensurate with their talent. What conclusions could be justified in regard to any gender link in this universal process, based on a sample of fewer than five women? And how can Loria be sure he knows their proper level when all the women up to the time of writing were systematically locked out of the best opportunities for professional advancement? Look at the twentieth century and see what becomes of Loria's argument that women never reach the top.

Finally, examine Loria's logic in the light of the cold facts of society: A woman who wished to have a career in mathematics would naturally be well advised to find a mentor with a well-established reputation, as Charlotte Angas Scott and Sof'ya Kovalevskaya did. A woman who did not do that would have no chance of being cited by Loria as an example, since she would never have been heard of. Is this argument not a classical example of catch-22?

Answer. This is too easy a shot. In fact, all the leading questions asked in the course of this problem give away the answer. Please enjoy writing your own rant against Loria.

4.10. Here is a policy question to consider. The primary undergraduate competition for mathematics majors is the Putnam Examination, administered the first weekend in December each year by the Mathematical Association of America. In addition to its rankings for the top teams and the top individuals, this examination also provides, for women who choose to enter, a prize for the highest-ranking woman. (The people grading the examinations do not know the identities of the entrants, and a woman can enter this competition without identifying herself to the graders.) Is this policy an important affirmative-action step to encourage talented young women in mathematical careers, or does it "send the wrong message," implying that women cannot compete with men on an equal basis in mathematics? If you consider it a good thing, how long should it be continued? Forever?

If not, what criterion should be used to determine when to discontinue the separate category? Bear in mind that the number of women taking the Putnam Examination is still considerably smaller than the number of men.

Answer. I'm on the side of "gender-neutrality" here and would prefer not to have a special category for women. It took some time for a woman (Emmy Noether) to come along who was in the very top rank of mathematicians of her generation, and it may be a while before a woman comes in among the top finishers on the Putnam Examination. But I'm very much against calling attention to the absence of women among the winners by giving them their own special category. It looks patronizing to me.

4.11. Continuing the topic of the Question 4.10, what criterion should be used to determine when affirmative action policies designed to overcome the effects of past discrimination against women will have achieved their aim? For example, are these policies to be continued until 50% of all mathematics professors are women within the universities of each ranking? (The American Mathematical Society divides institutions into different rankings according to the degrees they grant; there is also a less formal but still effective ranking in terms of the prestige of institutions.) What goal is being pursued: that each man and each woman should have equal access to the profession and equal opportunity for advancement in it, or that equal numbers of men and women will choose the profession and achieve advancement? Or is the goal different from both of these? If the goal is the first of these, how will we know when it has been achieved?

Answer. Unless we completely reorganize society from top to bottom and force people to make individual decisions about careers that will achieve a pre-ordained gender balance—a colossal project that very few people of either gender actually want at present—I believe it will be a long time before men and women in equal numbers choose to devote a major portion of their time to raising children and other domestic duties. In fact, it may never happen.[1] Given that, there is bound to be a statistical difference in the numbers of women and men working part-time and a statistical difference in the degree of commitment to a career. These differences will of course result in a statistical difference in men's and women's compensation. In addition, there *may* also be genuine differences in temperament that make some subjects and careers more attractive to members of one gender. (I do not assert that as fact, only as a possibility. Too little is known at present about objective gender differences; and those who assert that such differences do or do not exist are inevitably people who want the truth to be what they say it is.) Such differences, if they exist, are merely statistical and should not be thought to apply to any particular person. To clarify: I am *not* asserting that there is any gender-related difference between a male mathematician and a female mathematician. But there *may be* innate differences between men and women in general that result in more men than women wanting to be mathematicians or—what I think very likely in the near future—more women than men wanting to be mathematicians.

If the goal is to be that the whole set of women should have salaries and careers indistinguishable statistically from those of the whole set of men, I think it is entirely unrealistic. What is realistic (in my opinion) is that women and men performing comparable jobs with comparable experience should not be statistically distinguishable in their compensation and rank. Unfortunately, these statistical subtleties (not so subtle, really, but apparently beyond the comprehension of most news reporters) are usually lost in the relentless stream

[1] I am aware that many people think that no forcing of decisions would be necessary to bring about this outcome, that such an equal balance would inevitably result if society simply allowed each person a free choice. I think they are wrong.

of media reporting on the overall differences between men's and women's compensation. Where real discrimination against women still exists, I certainly want to see it rooted out. But one needs to locate the actual villains of the piece, not (as the news reporters so frequently do) merely report that "the average woman earns only 79 cents for every dollar the average man earns" as if that fact all by itself proved that someone was discriminating against women. The average man and the average woman are like the average book: they don't exist. In short, what I would like is to eliminate the identifiable sources of discrimination; I am against concluding that discrimination exists merely because of statistical differences.

Lest all this seem to be an anti-feminist rant, let me say that what I confidently expect to come about is a mathematics professoriate that is largely female. It will not happen for another generation, by which time I even more confidently expect to be dead. But my informal observation is that girls generally like school much better than boys and thereby get a better start (as a group, statistically) on nearly all careers. I think many parents who go to high school honors nights and see girls winning 90% of the honors, and (on the university level) women constituting 90% of the inductees to Phi Beta Kappa, will know what I mean. Unless we begin to pay more attention to the need for extra discipline that boys have, we will be essentially wasting their talents, as we wasted the talents of girls for many generations past. On the other hand, since only the most talented boys will become scientists and mathematicians, male scientists will be disproportionately represented at the top institutions. So much for my predictions. Time will tell.

CHAPTER 5

Counting

5.1. Find an example, different from those given in the text, in which English grammar makes a distinction between a set of two and a set of more than two objects.

Answer. Some grammarians will disagree with me, but I claim that *each other* and *one another* is such a distinction, as is the distinction between *between* and *among*.

5.2. Consider the following three-column list of number names in English and Russian. The first column contains the cardinal numbers (those used for counting), the second column the ordinal numbers (those used for ordering), and the third the fractional parts. Study and compare the three columns. The ordinal numbers and fractions and the numbers 1 and 2 are grammatically adjectives in Russian. They are given in the feminine form, since the fractions are always given that way in Russian, the noun *dolya*, meaning *part* or *share*, always being understood. If you know another language, prepare a similar table for that language, then describe your observations and inferences. What does the table suggest about the origin of counting?

English			**Russian**		
one	first	whole	odna	pervaya	tselaya
two	second	half	dve	vtoraya	polovina
three	third	third	tri	tret'ya	tret'
four	fourth	fourth	chetyre	chetvyortaya	chetvert'
five	fifth	fifth	pyat'	pyataya	pyataya
six	sixth	sixth	shest'	shestaya	shestaya

Answer. The corresponding table for French is as follows.

English			**French**		
one	first	whole	un	premier	entier
two	second	half	deux	seconde (deuxième)	demi
three	third	third	trois	troisième	tiers
four	fourth	fourth	quatre	quatrième	quart
five	fifth	fifth	cinq	cinquième	cinquième (partie)
six	sixth	sixth	six	sixième	sixième (partie)

For Japanese[1] it is as follows.

[1] Japanese grammar differs from English considerably, especially in regard to numbers. There are many ways of counting things in Japanese, depending on the nature of what is being counted.

English			Japanese		
one	first	whole	ichi	ichiban no	zenbu
two	second	half	ni	nibanme no	nibun no ichi
three	third	third	san	sanbanme no	sanbun no ichi
four	fourth	fourth	shi (yon)	yonbanme no	yonbun no ichi
five	fifth	fifth	go	gobanme no	gobun no ichi
six	sixth	sixth	roku	rokubanme no	rokubun no ichi

The French, English, and Russian languages show regularity setting in after the number 2. In Japanese the regularity applies even to the number 2. However, as noted, Japanese has more than one way of counting. The number names here come from Chinese and were superimposed on a Japanese counting system that already existed and is still used for counting certain things. In the other system the numbers 1, 2, 3, 4,... are *hitotsu*, *futatsu*, *mittsu*, *yottsu*,.... The old Japanese word for four is retained in many cases, since the Chinese word *shi* sounds like the Japanese word for death.

5.3. How do you account for the fact that the ancient Greeks used a system of counting and calculating that mirrored the notation found in Egypt, whereas in their astronomical measurements they borrowed the sexigesimal system of Mesopotamia? Why were they apparently blind to the computational advantages of the place-value system used in Mesopotamia?

Answer. It *may* simply be a case of which culture they encountered first. Once a system is established, transferring to a new system may require too much modification of an existing infrastructure. Why do we, for example, cling to the English system of spelling or the standard, but inefficient "qwertyuiop" keyboard on our computers?

5.4. A tropical year is the time elapsed between successive south-to-north crossings of the celestial equator by the Sun. A sidereal year is the time elapsed between two successive conjunctions of the Sun with a given star; that is, it is the time required for the Sun to make a full circuit of the ecliptic path that it appears (from Earth) to follow among the stars each year. Because the celestial equator is rotating (one revolution in 26,000 years) in the direction opposite to the Sun's motion along the ecliptic, a tropical year is about 20 minutes shorter than a sidereal year. Would you expect the flooding of the Nile to be synchronous with the tropical year or with the sidereal year? If the flooding is correlated with the tropical year, how long would it take for the heliacal rising of Sirius to be one day out of synchronicity with the Nile flood? If the two were synchronous 4000 years ago, how far apart would they be now, and would the flood occur later or earlier than the heliacal rising of Sirius?

Answer. Since the state of the Nile is determined by the weather and the climate and they are linked to the Sun, I would expect the flooding of the Nile to be linked to the tropical year. However, the flooding of the Nile was perhaps not so regular as some accounts would have us believe. Some historians think the "decline and fall" of the Old Kingdom late in the third millennium BCE may have been caused by the failure of the annual floods. The Nile flood begins in Egypt in early July, reaching a peak about two months later.

It takes about 72 years for the difference between tropical and sidereal years to amount to one day. That is, the heliacal rising of Sirius comes one day later every 72 years. In 4000 years, that would amount to about 8 weeks. Since the heliacal rising of Sirius is now at the end of July (Sirius is hidden by the sun during June and July), it would have occurred at the end of May in ancient times, giving the Egyptians a month to get ready for the flood.[2]

5.5. How many *Tzolkin* cycles are there in a Calendar Round?

Answer. 73 ($\frac{365 \cdot 52}{260}$).

5.6. The pattern of leap-year days in the Gregorian calendar has a 400-year cycle. Do the days of the week also recycle after 400 years?

Answer. Yes. A 400-year cycle has 97 leap years. Hence the total number of days in such a cycle is $400 \times 365 + 97 \equiv 1 \times 1 + 6 \equiv 0 \bmod 7$.

5.7. (*The revised Julian calendar*) The Gregorian calendar bears the name of the Pope who decreed that it should be used. It was therefore adopted early in many countries with a Catholic government, somewhat later in Anglican and Protestant countries. Countries that are largely Orthodox in faith resisted this reform until the year 1923, when a council suggested that century years should be leap years only when they leave a remainder of 2 or 6 when divided by 9. (This reform was not mandated, but was offered as a suggestion, pending universal agreement among all Christians on a date for Easter.) This modification would retain only two-ninths of the century years as leap years, instead of one-fourth, as in the Gregorian calendar. What is the average number of days in a year of this calendar? How does it compare with the actual length of a year? Is it more or less accurate than the Gregorian calendar?

Answer. In this calendar, the average number of days in a year, over a long period of time, will be

$$365 + \frac{1}{4} - \frac{7}{900} = \frac{164359}{450} \approx 365.24222.$$

That is as accurate as astronomy can get at the present time. However, with the newest methods of tracking time, we are now inserting "leap-seconds" occasionally to compensate for a very gradual slowing of the earth's rotation, so that the Gregorian calendar will not soon be out of synchrony with the seasons.

5.8. In constructing a calendar, we encounter the problem of measuring time. Measuring *space* is a comparatively straightforward task, based on the notion of congruent lengths. One can use a stick or a knotted rope stretched taut as a standard length and compare lengths or areas using it. Two lengths are congruent if each bears the same ratio to the standard length. In many cases one can move the objects around and bring them into coincidence. But what is meant by congruent *time intervals*? In what sense is the interval of time from 10:15 to 10:23 congruent to the time interval from 2:41 to 2:49?

Answer. The equality of time intervals as measured by standard clocks must be accepted as the starting point for quantitative science. It has no meaning outside of that context. The assumption is well justified by the consistency of the notion of equal time intervals when measured by different chronometers.

5.9. It seems clear that the decimal place-value system of writing integers is *potentially infinite*; that is there is no limit on the size of number that can be written in this system. But in practical terms, there is always a largest number for which a name exists. In ordinary language, we can talk about trillions, quadrillions, quintillions, sextillions, septillions, octillions, and so on. But somewhere before the number 10^{60} is reached, most people (except

[2] On the average. The month of warning leaves a reasonable safety margin in case the flood comes early.

Latin scholars) will run out of names. Some decades ago, a nephew of the American mathematician Edward Kasner (1878–1955) coined the name *googol* for the number 10^{100}, and later the name *googolplex* for $10^{10^{100}}$. This seems to be the largest number for which a name exists in English. Does there exist a positive integer for which no name *could* possibly be found, not merely an integer larger than all the integers that have been or will have been named before the human race becomes extinct? Give a logical argument in support of your answer. (And, while you are at it, consider what is meant by saying that an integer "exists.")

Answer. I think the answer here is "Yes," but I am willing to listen to contrary arguments if anybody has one. One such argument is that such an integer would have to be undescribable. If we had a description of such an integer, that description would be tantamount to a name. But since I think the structure of the numbers reflects some necessary structure of the universe and that it is meaningful to say that the objects in the universe and the relations among them are real, I see no reason not to use the colloquial modes of speech and say that there exist numbers no one will ever think of. Notice that English contains a potentially infinite set of words. We could rename the sequence of positive integers "ba," "baba," "bababa," "babababa," and so on, and then agree to use these names for the corresponding object only where a name does not already exist.

Here I am taking the philosophical position that integers "exist" in some sense—not the same sense in which physical objects exist. But I admit that assertions containing this use of the word *exist* are not to be taken in the same sense as the assertions that trees and shopping malls exist. What we mean when we say a number exists can be rephrased without using any word involving existence. For example, the assertion that the number three "exists" means that some sets consist of three members. The abstract concept of the number three and the notion of existence are merely linguistic conveniences, to help us in our quantitative reasoning. The elements of a three-element set (if they are physical) have the ordinary kind of existence. The set and the number three do not.

CHAPTER 6

Calculation

6.1. Double the hieroglyphic number $\begin{array}{ll} ||| & \cap \\ |||| & \cap\cap \end{array}$.

Answer. $\begin{array}{ll} || & \cap\cap\cap \\ || & \cap\cap\cap\cap \end{array}$.

6.2. Multiply 27 times 42 the Egyptian way.

Answer.

∗	1	42
∗	2	84
	4	168
∗	8	336
∗	16	672
Total	27	1134

6.3. (Stated in the Egyptian style.) Calculate with 13 so as to obtain 364.

Answer.

1	13	
2	26	
4	52	∗
8	104	∗
16	208	∗
28	364	Result

6.4. Problem 23 of the Ahmose Papyrus asks what parts must be added to the sum of $\overline{4}$, $\overline{8}$, $\overline{10}$, $\overline{30}$, and $\overline{45}$ to obtain $\overline{\overline{3}}$. See if you can obtain the author's answer of $\overline{9}\ \overline{40}$, starting with his technique of magnifying the first row by a factor of 45. Remember that $\frac{5}{8}$ must be expressed as $\overline{2}\ \overline{8}$.

Answer. This is one problem where the solution procedure is not shown in the Papyrus itself. Hence we can only conjecture the process used. When magnified by a factor of 45, the terms to be complemented total $23\ \overline{2}\ \overline{4}\ \overline{8}$, and the term to be reached is 30. What is to be added then is $6\ \overline{8}$. We need to divide this number by 45, that is, "calculate" with 45 so as to reach it. Apparently the scribe was observant enough to notice that $45 = 8 \cdot 5 + 1 \cdot 5$, so that the Horus-eye fractions could be used. Dividing 45 by 40 then gave $1\ \overline{8}$, so that only 5 was lacking, and it was produced by simply dividing 45 by 9. A different solution was suggested in the first edition of this book, and Gillings gives yet two more possibilities.

6.5. Problem 24 of the Ahmose Papyrus asks for a number that yields 19 when its seventh part is added to it, and concludes that one must perform on 7 the same operations that yield 19 when performed on 8. Now in Egyptian terms, 8 must be multiplied by $2\ \overline{4}\ \overline{8}$ in order to obtain 19. Multiply this number by 7 to obtain the scribe's answer, 16 $\overline{2}$ $\overline{8}$. Then multiply that result by $\overline{7}$, add the product to the result itself, and verify that you do obtain 19, as required.

Answer.

	1	7
*	2	14
	$\overline{2}$	3 $\overline{2}$
*	$\overline{4}$	1 $\overline{2}$ $\overline{4}$
*	$\overline{8}$	$\overline{2}$ $\overline{4}$ $\overline{8}$
Total	2 $\overline{4}$ $\overline{8}$	14 1 $\overline{2}$ $\overline{4}$ $\overline{2}$ $\overline{4}$ $\overline{8}$

This last number is 16 $\overline{2}$ $\overline{8}$. Multiplying by $\overline{7}$ yields (since the double of $\overline{7}$ is $\overline{4}$ $\overline{28}$) the number 2 $\overline{4}$ $\overline{14}$ $\overline{28}$ $\overline{56}$, and when we add this to 16 $\overline{2}$ $\overline{8}$ we get 18 $\overline{2}$ $\overline{4}$ $\overline{8}$ $\overline{14}$ $\overline{28}$ $\overline{56}$. The scribe had to recognize somehow that $\overline{14}$ $\overline{28}$ $\overline{56}$ is $\overline{8}$. Probably the Horus-eye representation $7 = 4 + 2 + 1$ was the secret of doing so, since $\overline{14}$, $\overline{28}$, and $\overline{56}$ are in those proportions.

6.6. Problem 33 of the Ahmose Papyrus asks for a quantity that yields 37 when increased by its two parts (two-thirds), its half, and its seventh part. Try to get the author's answer: The quantity is 16 $\overline{56}$ $\overline{679}$ $\overline{776}$. [*Hint*: Look in the table of doubles of parts for the double of $\overline{97}$. The scribe first tried the number 16 and found that the result of these operations applied to 16 fell short of 37 by the double of $\overline{42}$, which, as it happens, is exactly 1 $\overline{\overline{3}}$ $\overline{2}$ $\overline{7}$ times the double of $\overline{97}$.]

Answer. Using the "false position" method and trying 16, we find that the prescribed operations yield 36 $\overline{\overline{3}}$, $\overline{4}$ $\overline{28}$. This is very close to 37. To complement $\overline{\overline{3}}$, $\overline{4}$ $\overline{28}$ to 1, we multiply by 42, getting 28 10 $\overline{2}$ 1 $\overline{2}$, which is 40. We are therefore lacking 2 units and need to "calculate with 42 so as to obtain 2. In a way that seems weird to us, the author focused on the equation $97 = 42 + 28 + 21 + 6 = 42(1 + \frac{2}{3} + \frac{1}{2} + \frac{1}{7})$, which he would have thought of as the relation $\overline{42} = \overline{97}(1 + \frac{2}{3} + \frac{1}{2} + \frac{1}{7})$. Hence he had only to consult the table for the double of $\overline{97}$.

A more interesting question concerns the source of this problem. Where did that strange mixture of fractions in the statement of the problem arise? My guess is that it arose precisely from starting with the number 42, taking some integer "parts" of it, and adding them. The result happened to be 97. I believe the author started with the answer and tailored the problem to arrive at it. When the process is turned around and stated backwards, as in the Papyrus, the effect is to make the solution appear mysteriously, out of nowhere.

6.7. Verify that the solution to Problem 71 (2 $\overline{\overline{3}}$) is the correct *pesu* of the diluted beer discussed in the problem.

Answer. Since the original jug contained half a *hekat* of grain, it follows that one-eighth of a hekat was removed, leaving $\overline{4}$ $\overline{8}$ of a *hekat*. The reciprocal of this number is calculated in the ordinary way.

1	$\overline{4}$ $\overline{8}$
8	3
2 $\overline{\overline{3}}$	1

6.8. Compare the *pesu* problems in the Ahmose Papyrus with the following problem, which might have been taken from almost any algebra book written in the past century: *A radiator is filled with 16 quarts of a 10% alcohol solution. If it requires a 30% alcohol solution to protect the radiator from freezing when it is turned off, how much 95% solution must be added (after an equal amount of the 10% solution is drained off) to provide this protection?* Think of the alcohol as the grain in beer and the liquid in the radiator as the beer. The liquid has a *pesu* of 10. What is the *pesu* that it needs to have, and what is the *pesu* of the liquid that is to be used to achieve this result?

Answer. I actually took this problem from an algebra book. Unfortunately I do not remember exactly which algebra book; but that does not matter, since all algebra books are more or less interchangeable, and one cannot copyright a problem of this sort. The easiest way to compare the two problems is to solve the modern problem as we think an ancient Egyptian would have done. We are trying to mix a liquid with a *pesu* of 10 and a liquid with a *pesu* of 1 $\overline{19}$ to get 16 quarts of a liquid with a pesu of 3 $\overline{3}$. If we were Egyptians, we would first “calculate with 3 $\overline{3}$ so as to reach 16.” The result would be 4 $\overline{\overline{3}}$ $\overline{10}$ $\overline{30}$, which is the number of quarts of alcohol that must be in the final mixture. The amount now in it is 1 $\overline{2}$ $\overline{10}$ (found by dividing 16 by 10), so that we need to increase it by 3 $\overline{6}$ $\overline{30}$, which, as the scribe undoubtedly would have recognized, is 3 $\overline{5}$. Now each quart of *pesu*-10 solution contains $\overline{10}$ quarts of alcohol, and each quart of *pesu*-1 $\overline{19}$ solution contains $\overline{2}$ $\overline{4}$ $\overline{5}$ quarts of alcohol. Hence each quart replaced will increase the amount of alcohol by $\overline{2}$ $\overline{4}$ $\overline{10}$ quarts. We are thus looking for a number which, multiplied by $\overline{2}$ $\overline{4}$ $\overline{10}$ will yield 3 $\overline{5}$. Scaling the problem by multiplying it by 20 shows that we need to “calculate with 17 so as to obtain 64.” One can conjecture the following solution of that problem.

*	1	17
*	2	34
*	$\overline{\overline{3}}$	11 $\overline{3}$
*	$\overline{17}$	1
*	$\overline{34}$ $\overline{102}$	$\overline{\overline{3}}$

The answer would therefore be given as

$$3\ \overline{\overline{3}}\ \overline{17}\ \overline{34}\ \overline{102}.$$

Here the last two terms come from the table as the double of $\overline{51}$.

6.9. Verify that the solution $\overline{5}$ $\overline{10}$ given above for Problem 35 is correct, that is, multiply this number by 3 and by $\overline{3}$ and verify that the sum of the two results is 1.

Answer. This is straightforward:

*	1	$\overline{5}$ $\overline{10}$
*	2	$\overline{3}$ $\overline{15}$ $\overline{5}$
*	$\overline{3}$	$\overline{15}$ $\overline{30}$
Total	3 $\overline{3}$	$\overline{5}$ $\overline{10}$ $\overline{3}$ $\overline{15}$ $\overline{5}$ $\overline{15}$ $\overline{30}$

There is a lot of juggling of $\overline{5}$ and $\overline{15}$ when this last expression is condensed, since the double of $\overline{15}$ is $\overline{10}$ $\overline{30}$ and the double of $\overline{5}$ is $\overline{3}$ $\overline{15}$. In the end, one gets $\overline{\overline{3}}$ $\overline{5}$ $\overline{10}$ $\overline{30}$, which one can scale by 30 and see that the result is correct.

6.10. Why do you suppose that the author of the Ahmose Papyrus did not choose to say that the double of the thirteenth part is the seventh part plus the ninety-first part, that is,

$$\frac{2}{13} = \frac{1}{7} + \frac{1}{91}?$$

Why is the relation

$$\frac{2}{13} = \frac{1}{8} + \frac{1}{52} + \frac{1}{104}$$

made the basis for the tabular entry instead?

Answer. Undoubtedly the author preferred fractions with even numbers.

6.11. Generalizing Question 6.10, investigate the possibility of using the identity

$$\frac{2}{p} = \frac{1}{\left(\frac{p+1}{2}\right)} + \frac{1}{p\left(\frac{p+1}{2}\right)}$$

to express the double of the reciprocal of an odd number p as a sum of two reciprocals. Which of the entries in the table of Fig. 1 can be obtained from this pattern? Why was it not used to express $\frac{2}{15}$?

Answer. This pattern fits only the doubles of the reciprocals of 5, 7, 11, and 23 in the table. The difficulty seems to be that the second denominator tends to get large. With $\frac{2}{15}$, for example, it is 120, whereas the table gives the simpler decomposition $\frac{2}{15} = \frac{1}{10} + \frac{1}{30}$.

6.12. Why not simply write $\overline{13}\ \overline{13}$ to stand for what we call $\frac{2}{13}$? What is the reason for using two or three other "parts" instead of these two obvious parts?

Answer. This method leaves no way of combining the parts into simpler parts, for example replacing $\overline{6}\ \overline{30}$ with $\overline{5}$. Each time you double, you also double the number of terms you have to write down, and we all know how quickly doubling things leads to unwieldy amounts.

6.13. Could the ability to solve a problem such as Problem 35, discussed in Subsection 1.2 of this chapter, have been of any practical use? Try to think of a situation in which such a problem might arise.

Answer. One plausible conjecture is to calculate the volume of a new handmade pottery vessel in terms of a standard *hekat*. Thus if the vessel is filled three times and emptied into the standard *hekat* vessel, but does not quite fill the standard vessel, one could fill the new vessel approximately one-third the way full and discover that this amount just tops up the standard *hekat*. It would then be possible to use this computation to mark the volume of the vessel on its outside. (Please do not infer that the Egyptians actually *did* this. I have no direct evidence that they did. I am saying only that they *could have done it*.)

6.14. We would naturally solve many of the problems in the Ahmose Papyrus using an equation. Would it be appropriate to say that the Egyptians solved equations, or that they did algebra? What does the word *algebra* mean to you? How can you decide whether you are performing algebra or arithmetic?

Answer. They certainly found unknown numbers from properties that they must have. That is the essence of algebra. Their techniques were not ours, since they didn't write down and manipulate equations, but the underlying thought process was to reason from a description of a number to make it reveal itself in the notation they used for numbers. On that basis I say, "Yes, they did algebra."

6.15. Why did the Egyptians usually begin the process of division by multiplying by $\overline{\overline{3}}$ instead of the seemingly simpler $\overline{2}$?

Answer. As mentioned in the text, when you are putting numbers together from pieces, it saves labor to have the pieces as large as possible.

6.16. Early mathematicians must have been adept at thinking in terms of expressions. But considering the solutions to the riders-and-carts problem and the colorful language of Brahmagupta in relation to the Rule of Three, one might look at the situation from a different point of view. Perhaps these early mathematicians were good "dramatists." In any algorithm the objects we now call variables amount to special "roles" played, with different numbers being assigned to "act" in those roles; an algorithm amounts to the drama that results when these roles are acted. That is why it is so important that each part of the algorithm have its own name. The letters that we use for variables amount to names assigned to roles in the drama. A declaration of variables at the beginning of a program is analogous to the section that used to be titled "Dramatis Personæ" at the beginning of a play.

Explain long division from this point of view, using the roles of dividend, divisor, quotient, and remainder.

Answer. **Long Division: A Play in One Act** by Matthew Love.

Characters in the Drama:

Dividend. The cruel tyrant who has reduced *Quotient* to a mere cipher as the play begins. His pride goes before his fall, and in the end, he confesses himself vanquished and takes up the humble role of *Remainder*, trailing obediently behind *Divisor* and *Quotient*.

Divisor. The fearless champion who challenges *Dividend*, attacking him until the repeated assaults reduce *Dividend* to a mere shadow of his former self.

Quotient. Afraid to show himself at the opening of the play, this character grows in strength as the play proceeds.

Synopsis: The story is told as an allegory of a soccer match. As the play opens, *Dividend*, wearing a jersey with a large number on the front and back, comes to the front of the stage, proclaims himself the league champion, and boasts of his recent shutout victory over *Quotient*, who huddles miserably at stage left, his jersey bearing the number 0. *Divisor*, wearing a jersey with a smaller number than *Dividend*, then enters from stage right and challenges *Dividend* to a match. *Dividend* accepts the challenge; they clash, *Divisor* snatches *Dividend*'s jersey, and rips it off. *Dividend* is revealed to be wearing a second jersey underneath, with a number smaller than the original by exactly the number worn by *Divisor*. *Quotient* joyfully leaps up and rips off his own jersey, waving it in triumph and revealing a jersey underneath with the number 1 on front and back. *Dividend* boasts that he still has a bigger number than *Divisor*, and the two clash again. Each time they clash, *Divisor* rips off *Dividend*'s jersey, revealing a new jersey with a number that is smaller than its predecessor by an amount equal to *Divisor*'s jersey number. After each clash *Quotient* rips off his jersey and waves it, revealing another jersey underneath with a number one larger than its predecessor. Finally, *Dividend* is left with a jersey whose number is smaller than *Divisor*'s number. At that point, he renounces his league championship and the role of *Dividend*, accepting the humbler position of *Remainder*.

The one-act play of Matthew Love ends at this point. However, the Greek dramatist O. Euclid has written a more elaborate play, in which *Remainder* attempts to get revenge by attacking *Divisor*. He succeeds in his effort, and in the end usurps *Divisor*'s role, reducing him to the role of *Remainder*. Like all Greek drama, this cycle of revenge continues until ultimately *Divisor* and *Dividend/Remainder* annihilate each other. In still other versions of the drama neither of the two is ever able to score a permanent victory over the other, and the conflict continues forever.

6.17. Imitate the reasoning used in solving the problem of riders and carts above to solve Problem 17 of the *Sun Zi Suan Jing*. The problem asks how many guests were at a banquet if every two persons shared a bowl of rice, every three persons a bowl of soup, and every four persons a bowl of meat, leading to a total of 65 bowls. Don't use algebra, but try to explain the rather cryptic solution given by Sun Zi: Put down 65 bowls, multiply by 12 to obtain 780, and divide by 13 to get the answer.

Answer. We can see that 12 is the least common multiple of 2, 3, and 4. Each group of twelve people consumed six bowls of rice, four bowls of soup, and three bowls of meat, a total of 13 bowls. Since 65 bowls were used, it follows that there were five groups of twelve people, or 60 people. We obtained that answer by the "Rule of Three," dividing the number 65 by 13 and multiplying by 12. Sun Zi first multiplied by 12, then divided by 13.

6.18. Compare the following loosely interpreted problems from the *Jiu Zhang Suanshu* and the Ahmose Papyrus. First, from the *Jiu Zhang Suanshu*: Five officials went hunting and killed five deer. Their ranks entitle them to shares in the proportion $1:2:3:4:5$. What part of a deer does each receive?

Second, from the Ahmose Papyrus (Problem 40): 100 loaves of bread are to be divided among five people (in arithmetic progression), in such a way that the amount received by the last two (together) is one-seventh of the amount received by the first three (together). How much bread does each person receive?

Answer. We won't bother to solve these simple problems. Our interest here is in the kind of mathematics needed (proportional allocations in both cases). One wonders if the text of the problem reflects the real application of the mathematics. If we were talking about salaries, one can find echoes of this kind of problem in *Moby Dick*, when the sailors sign onto a whaling ship for a proportion of the ultimate profit, not knowing in advance, of course, how much that profit will be. The proportion is determined by negotiation, but the more skilled harpooneers were given a higher proportion than the mere rowers and deckhands. The two examples seem realistic, considering that they came from an economy in which coinage was probably not a universal means of exchange.

6.19. Compare the interest problem (Problem 20 of Chapter 3) from the *Jiu Zhang Suanshu* discussed above, with the following problem, taken from the American textbook *New Practical Arithmetic* by Benjamin Greenleaf (*1876*):

> The interest on $200 for 4 months being $4, what will be the interest on $590 for 1 year and 3 months?

Are there any significant differences at all in the nature of the two problems, written nearly 2000 years apart?

Answer. There is no significant difference at all. The arithmetic of finance is the same the world over.

6.20. Problem 4 in Chapter 6 of the *Jiu Zhang Suanshu* involves what is called *double false position*. The problem reads as follows: A number of families contribute equal amounts to purchase a herd of cattle. If the contribution (the same for each family) were such that seven families contribute a total of 190 [units of money], there would be a deficit of 330 [units of money]; but if the contribution were such that nine families contribute 270 [units of money], there would be a surplus of 30 [units of money]. Assuming that the families each contribute the correct amount, how much does the herd cost, and how many families are involved in the purchase? Explain the solution given by the author of the *Jiu Zhang Suanshu*, which goes as follows. Put down the proposed values (assessment to each family, that is, $\frac{190}{7}$ and $\frac{270}{9} = 30$), and below each put down the corresponding surplus or deficit (a positive number in each case). Cross-multiply and add the products to form the *shi* ($30 \cdot \frac{190}{7} + 330 \cdot 2709 = \frac{75000}{7}$). Add the surplus and deficit to form the *fa* ($330 + 30 = 360$). Subtract the smaller of the proposed values from the larger, to get the difference ($\frac{270}{9} - \frac{190}{7} = \frac{20}{7}$). Divide the *shi* by the difference to get the cost of the goods ($\frac{75000}{20} = 3750$); divide the *fa* by the difference to get the number of families ($\frac{360}{20/7} = 126$).

Answer. This is easy enough to *justify* if you merely write the problem down as the pair of linear equations

$$\begin{aligned} \frac{190}{7}x &= y - 330\,, \\ 30x &= y + 30\,. \end{aligned}$$

The question is not *justification* but *explanation*. How did the author know what to do? The operations suggested ought to have some relation to the commercial transactions imagined. The first cross-multiplication followed by addition can be thought of as attempting to make the first purchase 30 times, leading to a deficit of $30 \cdot 330$ and the second purchase 330 times, leading to a surplus of the same amount. Hence if one does both things, the amount of money paid will be exactly 360 (that is, 330 + 30) times the purchase price. The amount each family would have to contribute in order to make these 360 purchases is what the author called the *shi*. Hence it is only necessary to divide the *shi* by the *fa* to get the amount each family needs to contribute to make the purchase one time. However, the author doesn't do that. He probably reasoned that if the purchase was made according to the terms of the second transaction, and then reversed according to the terms of the first transaction, each family would have paid $\frac{20}{7}$ units of money and the whole village would be entitled to a refund of 360 units of money, which would represent the total amount of money paid. Hence the number of families must be $\frac{360}{20/7} = 126$.

6.21. Compare the pond-filling problem (Problem 26 of Chapter 6) of the *Jiu Zhang Suanshu* (discussed above) with the following problem from Greenleaf (*1876*, p. 125): *A cistern has three pipes; the first will fill it in 10 hours, the second in 15 hours, and the third in 16 hours. What time will it take them all to fill it?* Is there any real difference between the two problems?

Answer. Once again, there is no real difference in the two problems. Algebra is just as lamely seeking "practical" applications today as it was two thousand years ago.

6.22. The fair taxation problem from the *Jiu Zhang Suanshu* considered above treats distances and population with equal weight. That is, if the population of one county is double that of another, but that county is twice as far from the collection center, the two counties will have exactly the same tax assessment in grain and carts. Will this impose an equal burden on the taxpayers of the two counties? Is there a direct proportionality between distance and population that makes them interchangeable from the point of view of the taxpayers involved? Is the growing of extra grain to pay the tax fairly compensated by a shorter journey?

Answer. One would have to be a citizen of the place to know how relatively onerous paying a tax and traveling a distance over possibly rough terrain would be. This problem arises in all sorts of contexts, most especially in do-it-yourself projects. One must balance the value of one's own time and amateur skills against the budgetary impact of hiring a professional to do a job.

6.23. Perform the division $\frac{980}{45}$ following the method used by Brahmagupta.

Answer. Write $45 = 49 - 4$ We then have

$$\frac{980}{45} = \frac{980}{49}\Big(\frac{49}{45}\Big) = 20\Big(1 + \frac{4}{45}\Big) = 20 + \frac{4 \cdot 4}{9} = 21\frac{7}{9}\,.$$

6.24. Convert the sexagesimal number $5;\ 35,\ 10$ to decimal form and the number 314.7 to sexigesimal form.

Answer.

$$5;\ 35,\ 10 = 5+\frac{35}{60}+\frac{10}{3600} = 5+\frac{7}{12}+\frac{1}{360} = \frac{1800+210+1}{360} = \frac{2001}{360} = \frac{667}{120} = 5\frac{67}{120}\,.$$

$$314.7 = 314 + \frac{7}{10} = 314 + \frac{42}{60} = 314;\ 42\,.$$

6.25. As mentioned in connection with the lunisolar calendar, 19 solar years equal almost exactly 235 lunar months. (The difference is only about two hours.) In the Julian calendar, which has a leap year every fourth year, there is a natural 28-year cycle of calendars. The 28 years contain exactly seven leap-year days, giving a total of exactly 1461 weeks. These facts conjoin to provide a natural 532-year cycle ($532 = 28 \cdot 19$) of calendars incorporating the phases of the Moon. In particular, Easter, which is celebrated on the Sunday after the first full Moon of spring, has a 532-year cycle (spoiled only by the two-hour discrepancy between 19 years and 235 months). According to Simonov (*1999*), this 532-year cycle was known to Cyrus (Kirik) of Novgorod when he wrote his "Method by which one may determine the dates of all years" in the year 6644 from the creation of the world (1136 CE). Describe how you would create a table of dates of Easter that could, in principle, be used for all time, so that a user knowing the number of the current year could look in the table and determine the date of Easter for that year. How many rows and how many columns should such a table have, and how would it be used?

Answer. Well, the thing *could be* 1×532. Then you'd just have to divide the year by 532 and take the remainder to know the date of Easter. However, it would be simpler if it were 19×28. Then just divide the year by 19 and take the remainder to get the row and take the remainder on division by 28 to get the column. (Here I'd assume that the table user knew how to handle a remainder of 0. The old tables for Easter in prayer books used to make a special rule for that case in getting the "golden number," requiring the user to work with the integers from 1 to 19 instead of from 0 to 18. See the next problem for further explanation.)

6.26. From 1901 through 2099 the Gregorian calendar behaves like the Julian calendar, with a leap year every four years. Hence the 19-year lunar cycle and 28-year cycle of days interact in the same way during these two centuries. As an example, we calculate the date of Easter in the year 2039. The procedure is first to compute the remainder when 2039 is divided by 19. The result is 6 ($2039 = 19 \times 107 + 6$). This number tells us where the year 2039 occurs in the 19-year lunar cycle. In particular, by consulting the table below for year 6, we find that the first full Moon of spring in 2039 will occur on April 8. (Before people became familiar with the use of the number 0, it was customary to add 1 to this remainder, getting what is still known in prayer books as the *golden number*. Thus the golden number for the year 2039 is 7.)

We next determine by consulting the appropriate calendar in the 28-year cycle which day of the week April 8 will be. In fact, it will be a Friday in 2039, so that Easter will fall on April 10 in that year. The dates of the first full Moon in spring for the years of the lunar cycle are as follows. The year numbers are computed as above, by taking the remainder when the Gregorian year number is divided by 19.

Year	0	1	2	3	4	5	6
Full Moon	Apr. 14	Apr. 3	Mar. 23	Apr. 11	Mar. 31	Apr. 18	Apr. 8
Year	7	8	9	10	11	12	13
Full Moon	Mar. 28	Apr. 16	Apr. 5	Mar. 25	Apr. 13	Apr. 2	Mar. 22
Year	14	15	16	17	18		
Full Moon	Apr. 10	Mar. 30	Apr. 17	Apr. 7	Mar. 27		

Using this table, calculate the date of Easter for the years from 2040 through 2045. You can easily compute the day of the week for each of these dates in a given year, starting from the fact that March 21 in the year 2000 was a Tuesday. [*Note*: If the first full Moon of spring falls on a Sunday, Easter is the following Sunday.]

Answer. We can conveniently explain the answer by the following table. The only thing not in the table is the use of the year number modulo 28. It is easiest to keep track of the day name given to March 21 in these years. In 2040 that date will be a Wednesday. Hence it will be a Thursday in 2041, Friday in 2042, Saturday in 2043, Monday in 2044, and Tuesday in 2045. You can work out the day of the full moon from that point.

Year	**Year mod 19**	**Date of Full Moon**	**Date of Easter**
2040	7	March 28 (Wednesday)	April 1
2041	8	April 16 (Tuesday)	April 21
2042	9	April 5 (Saturday)	April 6
2043	10	March 25 (Wednesday)	March 29
2044	11	April 13 (Wednesday)	April 17
2045	12	April 2 (Sunday)	April 9

6.27. Prosthaphæresis can be carried out using only a table of cosines by making use of the formula

$$\cos\alpha\,\cos\beta = \frac{\cos(\alpha+\beta)+\cos(\alpha-\beta)}{2}.$$

Multiply 3562 by 4713 using this formula and a table of cosines. (It is fair to use your calculator as a table of cosines; just don't use its arithmetical capabilities.)

Answer. By using a calculator I find that $\alpha = \arccos(.3562) = 69.2329924597$ and $\beta = \arccos(.4713) = 61.8812845172$. Hence by prosthaphæresis,

$$3562 \times 4713 = 10^8 \cos(69.2329954597)\cos(61.881284572) =$$
$$= 10^8 \frac{\cos(131.014276977)+\cos(7.25170794244)}{2} = \frac{33575412}{2} = 16787706\,.$$

The intriguing thing about this computation is that one can use a trigonometric table in radians or degrees or any other unit of angle measure. The procedure yields the same (correct) result every way.

6.28. Do the multiplication $742518 \cdot 635942$ with pencil and paper without using a hand calculator, and time yourself. Also count the number of simple multiplications you do. Then get a calculator that will display 12 digits and do the same problem on it to see what errors you made, if any. (The author carried out the 36 multiplications and 63 additions in just under 5 minutes, but had two digits wrong in the answer as a result of incorrect carrying.)

Next, do the same problem using prosthaphæresis. (Again, you may use your hand calculator as a trigonometric table.) How much accuracy can you obtain this way? With a five-place table of cosines, using interpolation, the author found the two angles to be $50.52°$ and $42.05°$. The initial digits of the answer would thus be those of $\bigl(\cos(8.47°) + \cos(92.57°)\bigr)/2$, yielding 47213 as the initial digits of the 12-digit number. On the other hand, using a calculator that displays 14 digits, one finds the angles to be $50.510114088363°$ and $42.053645425939°$. That same calculator then returns all 12 digits of the correct answer as the numerical value of $\bigl(\cos(8.45646866242°) + \cos(92.563759514302°)\bigr)/2$. Compared with the time to do the problem in full the time saved was not significant.

Finally, do the problem using logarithms. Again, you may use your calculator to look up the logarithms, since a table is probably not readily available.

Answer. The author's results are already explained in the problem, except for the use of the logarithm. It should be done as follows, using natural logarithms.

$$\begin{aligned} \ln(742518) &= 13.5178023918 \\ \ln(635942) &= 13.3628626432 \\ \ln(742518 \times 635942) &= 26.880665035 \\ 742518 \times 635942 &= 472198381956\,. \end{aligned}$$

This answer is completely accurate. Tables would not have been so accurate unless they were very voluminous.

CHAPTER 7

Ancient Number Theory

7.1. Compute the sexagesimal representation of the number

$$\left(\frac{p/q - q/p}{2}\right)^2$$

for the following pairs of integers (p, q): $(12, 5)$, $(64, 27)$, $(75, 32)$, $(125, 54)$, and $(9, 5)$. Then correct column 1 of Plimpton 322 accordingly.

Answer. In sexagesimal notation rows 1, 2, 3, 4, and 15 should be respectively

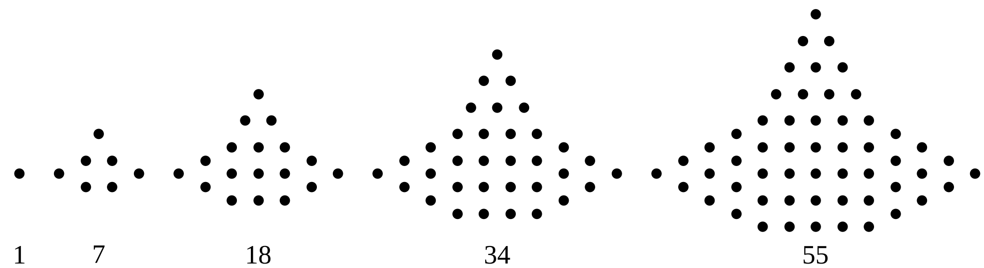
$$\begin{array}{l} 0;\ 59,\ 0,\ 15 \\ 0;\ 56,\ 56,\ 58,\ 14,\ 50,\ 6,\ 15 \\ 0;\ 55,\ 7,\ 41,\ 15,\ 33,\ 45 \\ 0;\ 53,\ 10,\ 29,\ 32,\ 52,\ 16 \\ 0;\ 23,\ 13,\ 46,\ 6,\ 40\,. \end{array}$$

7.2. On the surface the Euclidean algorithm looks easy to use, and indeed it *is* easy to use when applied to integers. The difficulty arises when it is applied to continuous objects (lengths, areas, volumes, weights). In order to execute a loop of this algorithm, you must be able to decide which element of the pair (a, b) is larger. But all judgments as to relative size run into the same difficulty that we encounter with calibrated measuring instruments: limited precision. There is a point at which one simply cannot say with certainty that the two quantities are either equal or unequal. Does this limitation have any practical significance? What is its theoretical significance? Show how it could give a wrong value for the greatest common measure even when the greatest common measure exists. How could it ever show that two quantities have *no* common measure?

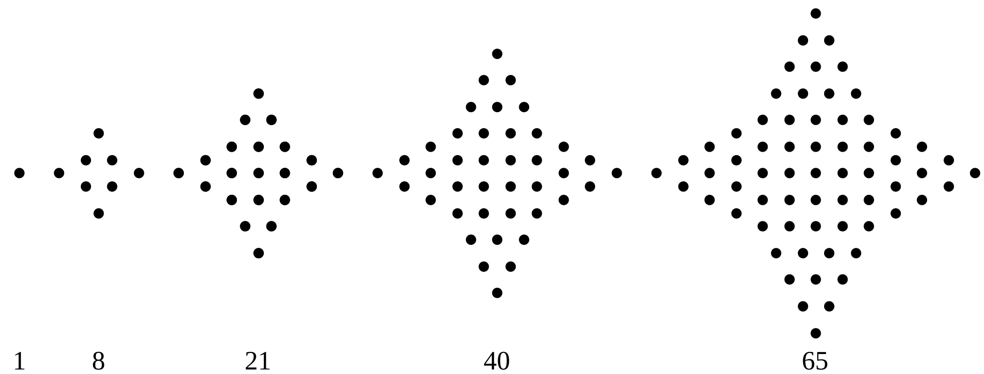
Answer. In both practical and theoretical terms, the limitation in precision simply means that any two quantities do have a common measure, namely the smallest observable quantity of the same type. Even if the greatest common measure really does exist, it will not have a *practical* unambiguous value, since it will be indistinguishable from an infinite number of other quantities. In practical terms, as noted above, the Euclidean algorithm cannot show that two *measured* quantities are incommensurable, since all measurements are given as integer multiples of the smallest unit.

7.3. The remainders in the Euclidean algorithm play an essential role in finding the greatest common divisor. The greatest common divisor of 488 and 24 is 8, so that the fraction $24/488$ can be reduced to $3/61$. The Euclidean algorithm generates two *quotients*, 20 and 3 (in order of generation). What is their relation to the two numbers? Observe the relation

$$\frac{1}{20 + \frac{1}{3}} = \frac{3}{61}$$



Heptagonal numbers ($\mathrm{Hp}_k = \frac{5}{2}k^2 - \frac{3}{2}k$)



Octagonal numbers ($\mathrm{O}_k = 3k^2 - 2k$)

If you find the greatest common divisor of 23 and 56 (which is 1) this way, you will generate the quotients 2, 2, 3, 3. Verify that

$$\frac{23}{56} = \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{3 + \cfrac{1}{3}}}}.$$

This expression is called the *continued fraction representation* representation of $23/56$. Formulate a general rule for finding the continued fraction representation of a proper fraction.

Answer. For a proper fraction $\frac{m}{n}$ (numerator smaller than denominator), if the successive quotients in the Euclidean algorithm are $q_1, q_2, \ldots, q_n$, then

$$\frac{m}{n} = \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cdots \cfrac{1}{q_n}}}.$$

7.4. Draw dot figures for the first five heptagonal and octagonal numbers. What kind of figure would you need for nonagonal numbers?

Answer. See the figure. You would need to glue 5 triangles onto a square to get a figure for nonagonal numbers. They wouldn't fit neatly.

7.5. Prove the formulas given in the caption of Fig. 1 for T_n, S_n, P_n, and H_n. Then prove that $S_n = T_n + T_{n-1}$, $P_n = S_n + T_{n-1} = T_n + 2T_{n-1}$, $H_n = P_n + T_{n-1} = T_n + 3T_{n-1}$. If $P_{k,n}$ is the nth k-gonal number, give a general formula for $P_{k,n}$ in terms of k and n.

Answer. The formula is $n^2 + (k-4)\frac{n(n-1)}{2}$, obtained by gluing $k-4$ triangles onto a square of side n. One can, of course, manipulate this expression to get $\frac{k-2}{2}n^2 - \frac{k-4}{2}n$

7.6. Prove that the Pythagorean procedure always produces a perfect number. That is, if $p = 2^n - 1$ is prime, then $N = 2^{n-1}p$ is perfect. This theorem is not difficult to prove nowadays, since the "parts" (proper divisors) of N are easy to list and sum.

Answer. The sum of the parts of $2^{n-1}(2^n - 1)$ is

$$(1 + 2 + \cdots + 2^{n-1}) + (2^n - 1)(1 + 2 + \cdots + 2^{n-2}) = 2^n(1 + \cdots + 2^{n-2}) + 2^{n-1} = \\ = 2^n(2^{n-1} - 1) + 2^{n-1} = 2^{n-1}(2^n - 2 + 1) = 2^{n-1}(2^n - 1)\,.$$

7.7. Let N_n be the nth perfect number, so that $N_1 = 6$, $N_2 = 28$, $N_3 = 496$, $N_4 = 8128$. Assuming that all perfect numbers are given by the Pythagorean formula, that is, they are of the form $2^{n-1}(2^n - 1)$ when $2^n - 1$ is a prime, prove that $N_{n+1} > 16N_n$ if $n > 1$. Conclude that there cannot be more than one k-digit perfect number for each k.

Answer. We'll assume $n > 2$. If $2^n - 1$ is prime, then certainly n is prime, since for $p > 1$ and $q > 1$ we find $2^{pq} - 1 = (2^p)^q - 1$, which is divisible by $2^p - 1$, which is larger than 1 and smaller than $2^{pq} - 1$. In particular, the next possible prime of the form $2^k - 1$ after $2^n - 1$ would be $2^{n+2} - 1$. Therefore the next possible perfect number after $2^{n-1}(2^n - 1)$ would be $2^{n+1}(2^{n+2} - 1)$. Since $2^{n+1} = 4 \cdot 2^{n-1}$ and $2^{n+2} - 1 > 2^{n+2} - 4 = 4 \cdot (2^n - 1)$, we see that the next *possible* perfect number after N_n is larger than $16N_n$. Hence certainly the next *actual* perfect number is larger than $16N_n$.

7.8. (*V. A. Lebesgue's proof of Euler's theorem on even perfect numbers*) Suppose that the perfect number N has the prime factorization $N = 2^\alpha p_1^{n_1} \cdots p_k^{n_k}$, where $p_1, \ldots, p_k$ are distinct odd primes and $\alpha, n_1, \ldots, n_k$ are nonnegative integers. Since N is perfect, the sum of *all* its divisors is $2N$. This means that

$$2^{\alpha+1} p_1^{n_1} \cdots p_k^{n_k} = (1 + 2 + \cdots + 2^\alpha)(1 + p_1 + \cdots + p_1^{n_1}) \cdots (1 + p_k + \cdots + p_k^{n_k}) \\ = \big(2^{\alpha+1} - 1\big)(1 + p_1 + \cdots + p_1^{n_1}) \cdots (1 + p_k + \cdots + p_k^{n_k})\,.$$

Rewrite this equation as follows:

$$\big(2^{\alpha+1} - 1\big) p_1^{n_1} \cdots p_k^{n_k} + p_1^{n_1} \cdots p_k^{n_k} = \\ = \big(2^{\alpha+1} - 1\big)(1 + p_1 + \cdots + p_1^{n_1}) \cdots (1 + p_k + \cdots + p_k^{n_k})\,,$$

$$p_1^{n_1} \cdots p_k^{n_k} + \frac{p_1^{n_1} \cdots p_k^{n_k}}{2^{\alpha+1} - 1} = (1 + p_1 + \cdots + p_1^{n_1}) \cdots (1 + p_k + \cdots + p_k^{n_k})\,.$$

Since the second term on the left must be an integer, it follows that $2^{\alpha+1} - 1$ must divide $p_1^{n_1} \cdots p_k^{n_k}$. This is not a significant statement if $\alpha = 0$ (N is an odd number). But if N is even, so that $\alpha > 0$, it implies that $2^{\alpha+1} - 1 = p_1^{m_1} \cdots p_k^{m_k}$ for integers $m_1 \le n_1, \ldots, m_k \le n_k$, *not all zero*. Thus, the left-hand side consists of the two *distinct* terms $p_1^{n_1} \cdots p_k^{n_k} + p_1^{r_1} \cdots p_k^{r_k}$. It follows that the right-hand side must also be equal to this sum. Now it is obvious that the right-hand side *contains* these two terms. That means the sum of the remaining terms on the right-hand side must be zero. But since the coefficients of all these terms are positive, there *can be* only two terms on the right. Since the right-hand side obviously contains $(n_1 + 1)(n_2 + 1) \cdots (n_k + 1)$ terms, we get the equation

$$2 = (n_1 + 1)(n_2 + 1) \cdots (n_k + 1).$$

Deduce from this equation that N must be of the form $2^{n-1}(2^n - 1)$ and that $2^n - 1$ is prime.

Answer. There really isn't much left to do. The only way that a product of positive integers can equal 2 is for one of the factors to be 2 and all the others to be 1. That means $n_1 = 1$ and $n_k = 0$ for $k > 1$. That is, $N = 2^\alpha p$, where p is a prime. But then the sum of the parts will be

$$(1 + 2 + \cdots + 2^\alpha) + p(1 + 2 + \cdots + 2^{\alpha-1}) = (2^\alpha - 1)(p + 1) + 2^\alpha = \\ = (2^\alpha - 1)p + 2^{\alpha+1} - 1 = N + 2^{\alpha+1} - p - 1\,.$$

Since this sum must equal N, the last equation implies that $p = 2^{\alpha+1} - 1$.

7.9. Generalize Diophantus' solution to the problem of finding a second representation of a number as the sum of two squares, using his example of $13 = 2^2 + 3^2$ and letting one of the numbers be $(\varsigma + 3)^2$ and the other $(k\varsigma - 2)^2$.

Answer. The equation $13 = (\varsigma+3)^2+(k\varsigma-2)^2$ is equivalent to $(k^2+1)\varsigma^2+(6-4k)\varsigma = 0$. That is, eliminating the case already found, in which $\varsigma = 0$,

$$\varsigma = \frac{4k-6}{k^2+1}.$$

Thus a general rational solution of $x^2 + y^2 = 13$ looks like

$$x = \frac{3k^2 + 4k - 3}{k^2 + 1}\,, \quad y = \frac{2k^2 - 6k - 2}{k^2 + 1}\,.$$

Here k can be any rational number whatsoever.

7.10. Take as a unit of time $T = \frac{1}{235}$ of a year, about 37 hours, 18 minutes, say a day and a half in close approximation. Then one average lunar month is $M = 19T$, and one average solar year is $Y = 235T$. Given that the Moon was full on June 1, 1996, what is the next year in which it will be full on June 4? Observe that June 4 in whatever year that is will be 3 days $(2T)$ plus an integer number of years. We are seeking integer numbers of months (x) and years (y), counting from June 1, 1996, such that $Mx = Yy + 2T$, that is (canceling T), $19x = 235y + 2$. Use the *kuttaka* to solve this problem and check your answer against an almanac. If you use this technique to answer this kind of question, you will get the correct answer most of the time. When the answer is wrong, it will be found that the full moon in the predicted year is a day earlier or a day later than the prescribed date. The occasional discrepancies occur because (1) the relation $M = 19T$ is not precise, (2) full moons occur at different times of day, and (3) the greatest-integer function is not continuous.

Answer. The *kuttaka* yields the quotients 12, 2, 1, 2. (The Euclidean algorithm would yield one more 2 if we carried it all the way to the end, but the *kuttaka* halts at the step *before* the Euclidean algorithm yields a zero remainder.) Since we have an even number of quotients, we put a positive 2 at the bottom, and our reduction algorithm then yields

$$\begin{matrix} 12 \\ 2 \\ 1 \\ 2 \\ 2 \\ 0 \end{matrix} \longrightarrow \begin{matrix} 12 \\ 2 \\ 1 \\ 4 \\ 2 \end{matrix} \longrightarrow \begin{matrix} 12 \\ 2 \\ 6 \\ 4 \end{matrix} \longrightarrow \begin{matrix} 12 \\ 16 \\ 6 \end{matrix} \longrightarrow \begin{matrix} 198 \\ 16 \end{matrix}.$$

Thus we get the solution $x = 198$, $y = 6$. Since y is the number of years, we see that the moon should have been full on June 4, 2002. A check of the almanac shows that it indeed was.

Comment. For the purposes of setting religious festivals, it really does not matter if the festival is a day late or a day early, so that this algorithm is suitable in all cases for fixing a calendar in advance.

7.11. Use Bhaskara's method to find two integers such that the square of their sum plus the cube of their sum equals twice the sum of their cubes. (This is a problem from Chapter 7 of the *Vija Ganita*.)

Answer. We make the problem determinate by fixing an arbitrary ratio: $y = kx$. In those terms, the equation $(x + y)^2 + (x + y)^3 = 2(x^3 + y^3)$ becomes

$$x = \frac{k + 1}{k^2 - 4k + 1}.$$

This expression is an integer when $k = 3$ ($x = -2$, $y = -6$), when $k = 4$ ($x = 5$, $y = 1$), and when $k = 5$ ($x = 1$, $y = 5$).

7.12. The Chinese mutual-subtraction algorithm (the Euclidean algorithm) can be used to convert a decimal expansion to a common fraction and to provide approximations to it with small denominators. Consider, for example, the number $e \approx 2.71828$. By the Euclidean algorithm, we get

$$\begin{aligned}
271,828 &= 2 \cdot 100,000 + 71,828 \\
100,000 &= 1 \cdot 71,828 + 28,172 \\
71,828 &= 2 \cdot 28,172 + 15,484 \\
28,172 &= 1 \cdot 15,484 + 12,688 \\
15,484 &= 1 \cdot 12,688 + 2,796 \\
12,688 &= 4 \cdot 2,796 + 1,504 \\
2,796 &= 1 \cdot 1,504 + 1,292 \\
1,504 &= 1 \cdot 1,292 + 212 \\
1,292 &= 6 \cdot 212 + 20 \\
212 &= 10 \cdot 20 + 12 \\
20 &= 1 \cdot 12 + 8 \\
12 &= 2 \cdot 8 + 4 \\
8 &= 2 \cdot 4
\end{aligned}$$

Thus the greatest common divisor of 271,828 and 100,000 is 4, and if it is divided out of all of these equations, the quotients remain the same. We can thus write

$$2.71828 = \frac{271828}{100000} = \frac{67957}{25000} = 2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{4 + \cdots}}}}}\,.$$

The first few partial fractions here give

$$\begin{aligned}
2 + \frac{1}{1} &= 3\,,\\
2 + \cfrac{1}{1 + \cfrac{1}{2}} &= 2\frac{2}{3} = \frac{8}{3} = 2.666\ldots\,,\\
2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1}}} &= \frac{11}{4} = 2.75\,,\\
2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{1}}}} &= 2\frac{5}{7} = \frac{19}{7} = 2.714285712485\ldots\,,\\
2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{4}}}}} &= 2 + \frac{23}{32} = \frac{87}{32} = 2.71875\,,
\end{aligned}$$

so that the approximations get better and better. Do the same with $\pi \approx 3.14159265$, and calculate the first five approximate fractions. Do you recognize any of these?

Answer. Let us take $\pi \approx 3.14159$. The Euclidean algorithm yields

$$\begin{aligned}
314,159 &= 3 \cdot 100,000 + 14,159\\
100,000 &= 7 \cdot 14,159 + 887\\
14159 &= 15 \cdot 887 + 854\\
887 &= 1 \cdot 854 + 33\\
854 &= 25 \cdot 33 + 29\\
33 &= 1 \cdot 29 + 4\\
29 &= 7 \cdot 4 + 1\\
4 &= 4 \cdot 1
\end{aligned}$$

The first few partial fractions are therefore

$$
\begin{aligned}
3+\frac{1}{7} &= \frac{22}{7} = 3.142857142857\ldots \\
3+\cfrac{1}{7+\cfrac{1}{15}} &= \frac{333}{106} = 3.141509433962\ldots \\
3+\cfrac{1}{7+\cfrac{1}{15+\cfrac{1}{1}}} &= \frac{355}{113} = 3.14159292035\ldots \\
3+\cfrac{1}{7+\cfrac{1}{15+\cfrac{1}{1+\cfrac{1}{25}}}} &= \frac{8793}{415} = 3.141589901058\ldots \\
3+\cfrac{1}{7+\cfrac{1}{15+\cfrac{1}{1+\cfrac{1}{25+\cfrac{1}{1}}}}} &= \frac{9563}{3044} = 3.14159001314\ldots
\end{aligned}
$$

At this point, we cannot really get any closer to the number π, since the approximation is now *better* than the decimal approximation we began with. This continued fraction, if carried on for two more steps, would get us the number 3.14159 *exactly*. Notice that these approximations go alternately above and below the number they are approximating (continued fractions always do that), but the last two are both less than π. It is interesting, however, that the two popular fractions $\frac{22}{7}$ and $\frac{355}{113}$ both occur in the sequence.

7.13. Can the pair of amicable numbers 1184 and 1210 be constructed from Thabit ibn-Qurra's formula?

Answer. No. Thus nothing like V. A. Lebesgue's theorem on perfect numbers will work for amicable numbers.

7.14. Solve the generalized problem stated by Matsunaga of finding an integer N that is simultaneously of the form $x^2 + a_1x + b_1$ and $y^2 + a_2y + b_2$. To do this, show that it is always possible to factor the number $(a_2^2 + 4b_1) - (a_1^2 + 4b_2)$ as a product mn, where m and n are either both even or both odd, and that the solution is found by taking $x = \frac{1}{2}\bigl(\frac{m-n}{2} - a_1\bigr)$, $y = \frac{1}{2}\bigl(\frac{m+n}{2} - a_2\bigr)$.

Answer. If the number $(a_2^2+4b_1)-(a_1^2+4b_2) = (a_2+a_1)(a_2-a_1)+4(b_1-b_2)$ is odd, that is, a_1 and a_2 have opposite parity, we can always take this number as m and then take $n = 1$ (and we have to do that if the number is prime). If it is even, then both $a_2 + a_1$ and $a_2 - a_1$ must be even (since they differ by the even number $2a_1$), and hence the number is a multiple of 4, so that it can be written as the product of two even numbers. The rest is merely a matter of verifying an identity.

7.15. Leonardo's solution to the problem of finding a second pair of squares having a given sum is explained in general terms, then illustrated with a special case. He considers the case $4^2 + 5^2 = 41$. He first finds two numbers (3 and 4) for which the sum of the squares *is* a square. He then forms the product of 41 and the sum of the squares of the latter pair, obtaining $25 \cdot 41 = 1025$. Then he finds two squares whose sum equals this number: 31^2 and 8^2 or 32^2 and 1^2. He thus obtains the results $\left(\frac{31}{5}\right)^2 + \left(\frac{8}{5}\right)^2 = 41$ and $\left(\frac{32}{5}\right)^2 + \left(\frac{1}{5}\right)^2 = 41$. Following this method, find another pair of rational numbers whose sum is 41. Why does the method work?

Answer. It works because of the identity $(a^2 + b^2)(c^2 + d^2) = (ad - bc)^2 + (ac + bd)^2$. If $c^2 + d^2 = g^2$, then

$$a^2 + b^2 = \left(\frac{ad - bc}{g}\right)^2 + \left(\frac{ac + bd}{g}\right)^2 .$$

If we take $c = 5$, $d = 12$, $g = 13$, for example, we get

$$41 = 4^2 + 5^2 = \left(\frac{23}{13}\right)^2 + \left(\frac{80}{13}\right)^2 .$$

Leonardo found two pairs of squares whose sum was 1025, and you can do the same by using the identity given above along with the identity $(a^2 + b^2)(c^2 + d^2) = (ad + bc)^2 + (ac - bd)^2$.

7.16. If the general term of the Fibonacci sequence is a_n, show that $a_n < a_{n+1} < 2a_n$, so that the ratio a_{n+1}/a_n always lies between 1 and 2. Assuming that this ratio has a limit, what is that limit?

Answer. Since it is obvious that the Fibonacci numbers are positive, it follows from the recursive relation that defines them that $a_{n+1} = a_n + a_{n-1} \leq 2a_n$. If L is the limit, it follows that L is between 1 and 2, and therefore, since

$$\frac{a_{n+1}}{a_n} = 1 + \frac{a_{n-1}}{a_n} ,$$

we have $L = 1 + \frac{1}{L}$. This means that

$$L^2 - L + 1 = 0 ,$$

and hence $L = \frac{1+\sqrt{5}}{2} = \Phi$.

7.17. Suppose that the pairs of rabbits begin to breed in the *first* month after they are born, but die after the second month (having produced two more pairs). What sequence of numbers results?

Answer. We need to keep track of two classes of rabbits, those that are newborn at the nth month (b_n), and those that are newborn at the $n-1$st month ($c_n = b_{n-1}$). The total number of rabbits a_n is $a_n = b_n + c_n = b_n + b_{n-1}$. Our recursion gives us $b_{n+1} = a_n$, so that b_n is a Fibonacci sequence, and hence a_n is just an advanced version of the Fibonacci sequence. In other words, where the sequence b_n is 1, 1, 2, 3, 5,..., the sequence a_n is 1, 2, 3, 5,....

7.18. Prove that if x, y, and z are relatively prime integers such that $x^2 + y^2 = z^2$, with x and z odd and y even, there exist integers u and v such that $x = u^2 - v^2$, $y = 2uv$, and $z = u^2 + v^2$. [*Hint*: Start from the fact that $x^2 = (z - y)(z + y)$, so that $z - y = a^2$ and $z + y = b^2$ for some a and b.]

Answer. Following the hint, we see that $z - y$ and $z + y$ are relatively prime odd integers whose product is x^2, and hence each must be a square. Let $z = a^2 + y = b^2 - y$, where a and b are odd integers and $x = ab$. Then $2y = b^2 - a^2 = (b - a)(b + a)$. It follows that $y = 2uv$ where $u = \frac{b+a}{2}$ and $v = \frac{b-a}{2}$. Then $u^2 - v^2 = ab = x$, and so $z^2 = (u^2 - v^2)^2 + 4u^2v^2 = (u^2 + v^2)$, as asserted.

CHAPTER 8

Numbers and Number Theory in Modern Mathematics

8.1. We know a mathematical algorithm for computing as many decimal digits of $\sqrt{2}$ as we have time for, and $\sqrt{2}$ has a precise representation in Euclidean geometry as the ratio of the diagonal of a square to its side. It is a provable theorem of Euclidean geometry that that ratio is the same for all squares, so that two observers using different squares should get the same result. To the extent that physical space really is Euclidean, this definition makes it possible to determine $\sqrt{2}$ empirically by measuring the sides and diagonals of physical squares. In that sense, we could theoretically determine $\sqrt{2}$ with arbitrarily prescribed precision by physical measurements. In particular, it makes perfectly good sense to ask what the 50th decimal digit of $\sqrt{2}$ is—it happens to be 4, but rounds up to 5—and we could try to get instruments precise enough to yield this result from measurement.

Consider, in contrast, the case of a physical constant, say the universal gravitational constant, usually denoted G_0, which occurs in Newton's law of gravitation:

$$F = G_0 \frac{Mm}{r^2}\,.$$

Here F is the force each of two bodies exerts on the other, M and m are the masses of the two bodies, and r is the distance between their centers of gravity. The accepted value of G, given as upper and lower assured limits, is $6.674215 \pm 0.000092\ \mathrm{N \cdot m^2/kg^2}$, although some recent measurements have cast doubt on this value. From a mathematical point of view, G_0 is determined by the equation

$$G_0 = \frac{Fr^2}{Mm}\,,$$

and its value is found—as Cavendish actually did—by putting two known masses M and m at a known distance r from each other and measuring the force each exerts on the other. The assertion that the ratio Fr^2/Mm is the same for all masses and all distances is precisely the content of *Newton's law of gravity*, so that two experimenters using different masses and different distances should get the same result. But Newton's law of gravity is not deducible from axioms; it is, rather, an empirical hypothesis, to be judged by its explanatory power and its consistency with observation. What should we conclude if two experimenters do *not* get the same result for the value of G_0? Did one of them do something wrong, or is Newton's law not applicable in all cases? Does it even *make sense* to ask what the 50th decimal digit of G_0 is?

Answer. The way the question is worded strongly suggests the answer. It makes sense to ask what the 50th decimal expansion of G_0 is provided the physical law that it comes from is accurate to that degree. We are not likely ever to know that, and in fact Newton's law of gravity has been replaced for very precise computations by Einstein's.

Another point that should be made in this connection concerns the relation between measurement and theory. The delicate measurements that are made using modern telescopes, microscopes, and other instruments are interpreted *on the basis of physical theories that explain how the instruments work*. If we are testing theory A by making instruments whose output is interpreted on the basis of theory B, then in effect we are testing A and B together. The apparent asymmetry that causes people to write as if only theory A is at risk comes about because in most cases we have more confidence in theory B. (Otherwise, we wouldn't use it to construct scientific equipment.)

8.2. You can represent $\sqrt{ab}$ geometrically by drawing putting a line of length b end-to-end with a line of length a, drawing a circle having this new line as diameter, and then drawing the perpendicular to the circle from the point where the two lines meet. To get $\sqrt{a}$ and $\sqrt{b}$, you would have to use Descartes' unit length I as one of the factors. Is it possible to prove by use of this construction that $\sqrt{ab}I = \sqrt{a}\sqrt{b}$? Was Dedekind justified in claiming that this identity had never been proved?

Answer. Before giving the answer, I should note that the argument about square roots was really a red herring. The Greeks could do a great many things with square roots that they couldn't do with cube roots. We are about to show how Euclid could have proved this identity. But Dedekind's point is no less cogent. It would have required a complicated induction (if it could be done at all) to show that the nth root is a multiplicative function.

I don't know of any place where it actually *was* proved. However, it *could have been* proved with rigor comparable to Euclid's, from the following considerations. Let A, B, C, I, and O be points in the plane such that OA stands for a number a, OB for a number B, OC for a number c (to be thought of as ab), and OI for a number to be thought of as 1. These interpretations are mapped into Euclidean geometry as the proportion $OI : OA = OB : OC$ ($c = ab$). Now let A', B', and C' be such that OA', OB', and OC' represent respectively $\sqrt{a}$, $\sqrt{b}$, and $\sqrt{c}$. That means $OI : OA' = OA' : OA$, $OI : OB' = OB' : OB$, and $OI : OC' = OC' : OC$. Finally, let D be a point such that the number d identified with OD represents the product $\sqrt{a}\sqrt{b}$. That is, $OI : OA' = OB' : OD$. What is to be shown is that OD equals (is congruent to) OC'. To do this, one needs to use certain propositions about proportional lines from Euclid's Book 6. In particular, his definition of composite ratio (see p. 292 of the textbook) implies that $(a : b).(c : d) = a : p$ provided $b : p = c : d$.[1] It follows that

$$(OA : OA').(OA' : OD) = OA : OD\,.$$

Likewise,

$$(OD : OB').(OB' : OB) = OD : OB\,.$$

However, since $OI : OA' = OA' : OA$ and $OI : OA' = OB' : OD$, it follows that $OA' : OA = OB' : OD$, and by a trivial application of the definition of proportion,

$$OA : OA' = OD : OB'\,.$$

By symmetry (or by simply repeating the argument),

$$OA' : OD = OB' : OB\,.$$

[1] Euclid uses this definition mostly to show in Proposition 23 of Book 6 that the ratio of two rectangles is the composite of the ratio of their heights and the ratio of their widths and to show that similar polygons are proportional to the duplicate ratios of their sides. (We would say the squares of their sides.)

Then by substitution of equals (which is allowed by Proposition 2 of Euclid's *Data*), we get

$$OA : OD = (OA : OA').(OA' : OD) = (OD : OB').(OB' : OB) = OD : OB\,.$$

That is, $OA : OD = OD : OB$. According to Proposition 16 of Book 6, this means the square on OD equals the rectangle on OA and OB. That is, OD which by definition represents the product $\sqrt{a}\sqrt{b}$, actually represents $\sqrt{ab}$, which was to be proved.

8.3. Try to give a definition of real numbers—perhaps using decimal expansions—that will enable you to say what the numbers $\sqrt{2}$, $\sqrt{3}$, and $\sqrt{6}$ are, and how they can be added and multiplied. Does your definition enable you to prove that $\sqrt{2}\sqrt{3} = \sqrt{6}$?

Answer. We shall assume that the numbers to be added are known only with *potentially* infinite precision. Our model for this is the way we express $\sqrt{2}$ decimally. Although we will never at any finite point in time know all of the decimal digits of this number, we do have a systematic procedure for finding initial segments of it that are arbitrarily long. What is (in my view) the hardest thing to accept about Dedekind's definition of a real number is that it seems to require knowledge with *actually infinite precision*; and the same problem affects other variant definitions, such as Weierstrass' definition in terms of Cauchy sequences of rational numbers. This problem seems to be unavoidable. (At least, I don't know any way to avoid it.) The following scenario is intended only to push the problem as far away as possible.

It will simplify things if we regard real numbers as binary expansions; that is, an integer (given as a finite binary expansion) together with a fraction between 0 and 1 (given as an infinite sequence of zeros and ones) in which we require that there be an infinite number of zeros, just to make the expansions unique. Since adding integers is a finite process, we concentrate on how to add two fractional parts. Again, we assume that such a fractional part is known if we can find arbitrarily long initial segments of it. Our job is to show how to add two such fractional parts $A = [a_1, a_2, \ldots, a_n, \ldots]$ and $B = [b_1, b_2, \ldots, b_n, \ldots]$ so as to get $C = [c_1, c_2, \ldots, c_n, \ldots] = A + B$. In other words, by our interpretation of what it means to define such a number, we have to show how it is possible to get arbitrarily long initial segments of C, knowing arbitrarily long initial segments of A and B and to assure that C contains an infinite number of zeros. We note that if A has only a finite number of ones in its expansion, then the expansion of C coincides with the expansion of B from some point on, and the problem is trivial. Hence we assume that the expansions of both A and B contain an infinite number of ones.

The procedure we have in mind is best illustrated with an example. Suppose $A = [0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 1, 0, 0, \ldots]$ and $B = [1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 1, \ldots]$. We compare the two expansions until we come to the first place where both expansions have the same digit. In this example, that occurs in the seventh place and the matching digit is a one. We then write $A = [0, 1, 1, 0, 0, 0, 1] + A_1$, where $0 < A_1 < 2^{-7}$—the second inequality is strict because of the assumption that the expansion of A contains an infinite number of zeros—and $B = [1, 0, 0, 1, 1, 1, 1] + B_1$, where $0 < B_1 < 2^{-7}$. We then add the parts separately, getting $C = 1 + [0, 0, 0, 0, 0, 0, 0] + A_1 + B_1$. We now see that the first six digits of C must be zero. For $A_1 + B_1 < 2^{-6}$, so that this sum could not require any ones before the seventh binary place. Hence we know the first six digits are zeros. The seventh is temporarily a zero. It will become a one if and only if the next match is also a pair of ones. Since these expansions both have infinitely many zeros, we will not have a pair of ones in every place, and hence, eventually, C will get a permanent zero among its digits.

The problem now reduces to adding $A_1 = [0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 0, \ldots]$ and $B_1 = [0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, \ldots]$, and the first match after the initial string of seven zeros occurs in the thirteenth place, where each has a zero. We therefore write

$$A_1 = [0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 0] + A_2\ ,$$

where $0 < A_2 < 2^{-13}$, and

$$B_1 = [0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0] + B_2\ ,$$

where $0 < B_2 < 2^{-13}$, and set $C_1 = [0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 0] + A_2 + B_2$, where $A_2 + B_2 < 2^{-12}$. Since the twelfth binary digit of $A_2 + B_2$ will be zero, we now know the first twelve digits of C_1, and since the first seven digits are zero, we also know the first twelve digits of C. Observe that this match leaves the zero that was in the seventh place intact. That is true in general, each time the match switches from a pair of ones to a pair of zeros, the zero inserted at the place of the previous match gets "frozen" there permanently. On the other hand, each matching pair of 1's at the Nth place inserts a permanent 0 at place $N - 1$, unless the previous match occurred at place $N - 1$. Thus, if there are infinitely many gaps between matching places, there will be infinitely many zeros in the number constructed as the sum, and if there are not, then the two numbers are identical from some point on, and the sum from that point on is a simple matter of shifting one of the expansions leftward by one place. In either case, there will be an infinite number of zeros. Finally, if there are no more matches from some point on, then the expansions are complementary from that point on, and it is only necessary to add insert a one in the last place where a match occurred (temporarily containing a zero because of the match) and set all remaining digits equal to zero. Thus we are assured of getting a binary expansion of the required form for C.

This process can be continued as long as necessary. It shows that only one problem arises in defining addition of real numbers via binary expansions: We must be able to tell if two binary expansions have only a finite number of matching digits. This problem brings back the difficulty of having to know infinitely much information in order to define addition. I do not see how it can be overcome. It is worth noting, however, that the set of pairs for which there are only finitely many matching digits have a sum that is a finite binary expansion. All such pairs therefore lie on a countable set of lines in the plane, and hence form a set of measure zero. In other words, it is infinitely unlikely that randomly chosen numbers A and B will exhibit this difficulty.

Once addition has been defined, multiplication is a fairly straightforward matter, since it amounts to a sequence of register shifts followed by addition, which has already been defined.

The square root algorithm works very simply in binary notation, since the doubling procedure that it involves is merely a matter of adding a zero at the end of each number to be doubled. If we know $2n$ binary digits of x, we can find n binary digits of $\sqrt{x}$ by this procedure. Hence $\sqrt{x}$ can be regarded as defined by our interpretation of what a number is, provided x itself is defined.

Because of the way the algorithm works, it is readily apparent that $\sqrt{4x} = 2\sqrt{x}$ (since multiplying by 4 merely moves the binary point marker two digits to the left). By repeated application of this principle, one can reduce the problem of showing $\sqrt{ab} = \sqrt{a}\sqrt{b}$ to the case when a and b lie between 0 and 1. By working with inequalities, for any two numbers x, x', y, y' less than 1, one can show that $|(xy)^2 - (x'y')^2| \leq \bigl|\bigl((xy) + (x'y')\bigr)\bigl((xy) - (x'y')\bigr)\bigr| \leq 2|(x - x')y + (y - y')x'| \leq 2\bigl(|x - x'| + |y - y'|\bigr)$. Thus if x' is sufficiently close to x and y' is sufficiently close to y, then $x'y'$ is close to xy. Applying this in a particular case where x' and y' are finite binary numbers approximating $x = \sqrt{a}$ and $y = \sqrt{b}$, we find that $(x'y')^2$ is a close approximation to $(\sqrt{a}\sqrt{b})^2$. But for finite binary numbers we know that $(x'y')^2 = (x')^2(y')^2$, which is a good approximation to x^2y^2, that is, to $(\sqrt{a})^2(\sqrt{b})^2 = ab$. Thus, in this interpretation of real numbers, we can prove that $\sqrt{ab} = \sqrt{a}\sqrt{b}$, and we could do the same for nth roots.

8.4. Use the method of infinite descent to prove that $\sqrt{3}$ is irrational. [*Hint*: Assuming that $m^2 = 3n^2$, where m and n are positive integers having no common factor, that is, they are as small as possible, verify that $(m-3n)^2 = 3(m-n)^2$. Note that $m < 2n$ and hence $m - n < n$, contradicting the minimality of the original m and n.]

Answer. Notice that if we had positive integers m and n satisfying this equation, there would have to be a smallest positive integer m for which there exists an n such that the equation holds. That n would then have to be the smallest positive integer n for which there exists an m such that the equation holds, since m and n are strictly increasing (one-to-one) functions of each other. If $m \geq 2n$, then $m^2 \geq 4n^2 > 3n^2$, so we certainly have $m < 2n$, and hence $0 < m - n < n$. But the equation implies that $(3n-m)^2 = 3\cdot(3n^2) - 6mn + m^2 = 3m^2 - 6mn + 3n^2 = 3(m-n)^2$. The assertion that m and n have no common factor is of course true if they are minimal integers. But since no such integers m and n even exist, it wasn't really relevant to include that property in the statement of the problem.

8.5. Show that $\sqrt[3]{3}$ is irrational by assuming that $m^3 = 3n^3$ with m and n positive integers having no common factor. [*Hint*: Show that $(m-n)(m^2 + mn + n^2) = 2n^3$. Hence, if p is a prime factor of n, then p divides either $m - n$ or $m^2 + mn + n^2$. In either case p must divide m. Since m and n have no common factor, it follows that $n = 1$.]

Answer. The work is essentially already done. Once we get to the point that $n = 1$, we are saying that if $\sqrt[3]{3}$ is rational, it must be an *integer*. But certainly there is no integer whose cube equals 3.

8.6. Suppose that x, y, and z are positive integers, no two of which have a common factor, none of which is divisible by 3, and such that $x^3 + y^3 = z^3$. Show that there exist integers p, q, and r such that $z - x = p^3$, $z - y = q^3$, and $x + y = r^3$. Then, letting $m = r^3 - (p^3 + q^3)$ and $n = 2pqr$, verify from the original equation that $m^3 = 3n^3$, which by Problem 8.5 is impossible if m and n are nonzero. Hence $n = 0$, which means that $p = 0$ or $q = 0$ or $r = 0$, that is, at least one of x and y equals 0. Conclude that no such positive integers x, y, and z can exist.

Answer. We give the proof that $z - x = p^3$. The proof that q and r exist is the same. We have

$$(z-x)(z^2 + xz + x^2) = z^3 - x^3 = y^3 .$$

But $z - x$ and $z^2 + xz + x^2$ must be relatively prime. For if n is a prime number that divides both, then n divides $(z-x)^2 = z^2 - 2xz + x^2$, and hence n divides $3xz$. But n is not 3, since if $z - x$ is divisible by 3, so is y, contrary to hypothesis. Hence n divides either x or z and so (since it divides $z - x$) both of them, which is again contrary to hypothesis. Now we have two relatively prime numbers whose product is a cube, and hence each must be a cube. That is, $z - x = p^3$ for some integer p, as asserted. The rest of the argument is merely a very messy computation.

8.7. Verify that

$$27^5 + 84^5 + 110^5 + 133^5 = 144^5.$$

[See L. J. Lander and T. R. Parkin, "Counterexample to Euler's conjecture on sums of like powers," *Bulletin of the American Mathematical Society*, **72** (1966), p. 1079. Smaller counterexamples to this conjecture have been discovered more recently.]

Answer. I will assume the reader can compute these numbers accurately, or is willing to trust a calculator to do it.

8.8. Prove Fermat's little theorem by induction on a. [*Hint*: The theorem can be restated as the assertion that p divides $a^p - a$ for every positive integer a. Use the binomial theorem to show that $(a+1)^p - (a+1) = mp + a^p - a$ for some integer m.]

Answer. The assertion in the hint amounts to the statement that a prime number p divides all the binomial coefficients $C_{p,k} = \frac{p!}{k!(p-k)!}$ for $k = 1, 2, \ldots, p-1$. But that is easy, since $C_{p,k}$ is known to be an integer. For

$$k!(p-k)!C_{p,k} = p!\,.$$

Hence p divides the left-hand side here. Hence it must divide either $C_{p,k}$ or some integer smaller than the maximum of k and $p-k$. The latter is impossible, since p is larger than all these integers. It follows that p divides $C_{p,k}$.

8.9. Verify the law of quadratic reciprocity for the primes 17 and 23 and for 67 and 71.

Answer. Since $17 = 4 \cdot 4 + 1$, we observe that the squares modulo 17 are 0, 1, 4, 9, 16, 8, 2, 15, 13, and that 6 is not among them. It must therefore be true that 17 is not a square modulo 23, and indeed that is true: The squares modulo 23 are 0, 1, 4, 9, 16, 2, 13, 3, 18, 12, 8, 6, and 17 is not among them.

Since both 67 and 71 are equal to 3 modulo 4, we expect precisely one to be a square modulo the other. Since $71 \equiv 4 = 2^2 \bmod 67$, we conclude that 67 is not among the 36 squares modulo 71. These squares are 0, 1, 4, 9, 16, 25, 36, 49, 64, 10, 29, 50, 2, 27, 54, 12, 43, 5, 40, 6, 45, 15, 58, 32, 8, 57, 37, 19, 3, 60, 48, 38, 30, 24, 20, 18, and indeed 67 is not among them.

8.10. Show that the factorization of numbers of the form $m + n\sqrt{-3}$ is *not* unique by finding two different factorizations of 4. Is factorization unique for numbers of the form $m + n\sqrt{-2}$?

Answer. We are helped in this effort by the existence of the valuation $V(m + n\sqrt{-3}) = m^2 + 3n^2$. This valuation has the property that $V(zw) = V(z)V(w)$. In particular, since $V(z) \ge 1$ for all nonzero z, we see that it is impossible to factor 2, except trivially as $2 = \pm 1 \cdot \pm 2$. (If $2 = zw$, then $V(z)V(w) = V(2) = 4$. That leaves only the possibility $V(z) = 1$, $V(w) = 4$ or $V(z) = 4$, $V(w) = 1$, since $V(z) = 2 = V(w)$ is impossible. (The equation $m^2 + 3n^2 = 2$ has no integer solutions.) Hence either $z = \pm 1$ or $w = \pm 1$. Likewise, it is impossible to factor $1 + \sqrt{-3}$, except trivially, and for the same reason. Thus $1 \pm \sqrt{-3}$ and 2 are irreducible numbers in this system. But

$$2 \cdot 2 = 4 = (1 + \sqrt{-3})(1 - \sqrt{-3})\,.$$

Factorization *is* unique for numbers of the form $m + n\sqrt{-2}$, and that is because a Euclidean algorithm exists for this structure: For any two elements of it, say z and $w \ne 0$, there exist a quotient q and a remainder r such that

$$z = qw + r$$

and $V(r) < V(w)$. Here, of course, $V(m + n\sqrt{-2}) = m^2 + 2n^2$.

To see why this is true, let $z = m + n\sqrt{-2}$ and $w = s + t\sqrt{-2}$. If $q = x + y\sqrt{-2}$, then $qw = (sx - 2ty) + (sy + tx)\sqrt{-2}$. We claim it is possible to choose the integers x and y such that $(sx - 2ty - m)^2 + 2(sy + tx - n)^2 < s^2 + 2t^2$. In fact, the exact solution of the equations

$$\begin{aligned} sx^* - 2ty^* &= m\,, \\ tx^* + sy^* &= n\,, \end{aligned}$$

is

$$\begin{aligned} x^* &= \frac{ms - 2tn}{s^2 + 2t^2}\,, \\ y^* &= \frac{2n - mt}{s^2 + 2t^2}\,. \end{aligned}$$

Of course, this choice of x^* and y^* may be nonintegers. Let x and y be the nearest integers to x^* and y^* respectively. Then $|x - x^*| \le \frac{1}{2}$ and $|y - y^*| \le \frac{1}{2}$, and so we find that

$$\begin{aligned} (sx - 2ty - m)^2 &= \bigl(s(x - x^*) - 2t(y - y^*)\bigr)^2 \\ &= s^2(x - x^*)^2 - 4st(x - x^*)(y - y^*) + 4t^2(y - y^*)^2\,, \\ (sy + tx - n)^2 &= \bigl(s(y - y^*) + t(x - x^*)\bigr)^2 \\ &= s^2(y - y^*)^2 + 2st(x - x^*)(y - y^*) + t^2(x - x^*)^2\,. \end{aligned}$$

It follows that

$$\begin{aligned} (sx - 2ty - m)^2 + 2(sy + tx - n)^2 &= s^2\bigl(x - x^*)^2 + 2(y - y^*)^2\bigr) + \\ &\quad + 2t^2\bigl((x - x^*)^2 + 2(y - y^*)^2\bigr) \le \frac{3}{4}\bigl(s^2 + 2t^2\bigr) < s^2 + 2t^2\,. \end{aligned}$$

(Notice that this argument breaks down if $\sqrt{-2}$ is replaced by $\sqrt{-3}$.) From this crucial result, it follows that a division algorithm exists for finding a greatest common divisor for z and w. Since the valuation of the remainder r decreases at each stage, it forms a strictly decreasing sequence of nonnegative integers until it reaches the value 0, which it must do in a finite number of steps. But when the remainder is 0, the division comes out even, and so the last non-zero remainder is the greatest common divisor. The same argument, based on a chain of equations of the form $z = qw + r$, shows that the greatest common divisor of z and w is a multiple of any other common divisor, and hence if two different divisors result from possibly different choices of the remainder at each stage in this algorithm, they must divide each other and therefore must be either equal or negatives of each other. That means, in particular, that if an irreducible number p divides a product ab in this system, it must divide either a or b. For, if not, its greatest common divisors p_a and p_b with both a and b respectively have the property that p_a divides a and p, p_b divides b and p. Since $p = \pm p_a$ or $p = \pm p_b$, that means p divides a or p divides b. From that crucial fact, it follows that a factorization into irreducible elements in this system is unique, up to factors of ± 1.

8.11. Prove that the number of primes less than or equal to N is at least $\log_2(N/3)$, by proceeding as follows. Let $p_1, \ldots, p_n$ be the prime numbers among $1, \ldots, N$, and let $\theta(N)$ be the number of square-free integers among $1, \ldots, N$, that is, the integers not divisible by any square number. We then have the following relation, since it is known that

$\sum_{k=1}^{\infty}(1/k^2) = \pi^2/6$.

$$\begin{aligned}\theta(N) &> N - \sum_{k=1}^{n}\left[\frac{N}{p_k^2}\right] \\ &> N\left(1 - \sum_{k=1}^{n}\frac{1}{p_k^2}\right) \\ &> N\left(1 - \sum_{k=2}^{\infty}\frac{1}{k^2}\right) \\ &= N\left(2 - \frac{\pi^2}{6}\right) > \frac{N}{3}.\end{aligned}$$

(Here the square brackets denote the greatest-integer function.) Now a square-free integer k between 1 and N is of the form $k = p_1^{e_1}\cdots p_n^{e_n}$, where each e_j is either 0 or 1. Hence $\theta(N) \le 2^n$, and so $n > \log_2(N/3)$. This interesting bit of mathematical trivia is due to the Russian–American mathematician Joseph Perott (1854–1924).

Answer. We will assume that $N > 36$, just for convenience. You can do the other 35 cases yourself. In the first inequality here, we have subtracted the number of integers not greater than N that are divisible by p_k^2. If we do this separately for each N, we get something strictly smaller than $\theta(N)$, since some integers (36, for example) are divisible by the squares of two different primes. The second inequality is merely the fact that $[k] \le k$ for positive integers k. The third is the fact that all primes are at least 2 and that not every integer is a prime. The last equality uses the stated value for the sum of the series and the fact that $\pi^2 < 10$.

8.12. Assuming that $\lim_{n\to\infty}\dfrac{\log(n)\pi(n)}{n}$ exists, use Chebyshëv's estimates to show that this limit must be 1 and hence that Legendre's estimate cannot be valid beyond the first term.

Answer. Please note that the logarithm is a natural logarithm, which should probably have been denoted ln for the audience that is expected to read this book. According to one of the estimates, with $\alpha = 1$, $m = 2$, we find

$$\frac{\ln(n)\pi(n)}{n} > \frac{\ln(n)\int_2^n \frac{1}{\ln x}\,dx}{n} - \frac{1}{\ln(n)}\,.$$

We can evaluate the limit of the right-hand side using l'Hôpital's rule. It is

$$1 + \lim_{n\to\infty}\frac{1}{n}\int_2^n \frac{1}{\ln(x)}\,dx\,.$$

Another application of l'Hôpital's rule shows that the limit is

$$1 + \lim_{n\to\infty}\frac{1}{\ln(n)} = 1\,.$$

Hence if

$$\lim_{n\to\infty}\frac{\ln(n)\pi(n)}{n}$$

exists, it cannot be less than 1. A similar argument with the other inequality shows that this limit cannot be less than 1.

CHAPTER 9

Measurement

9.1. Show how it is possible to square the circle using ruler and compass given the assumption that $\pi = (16\sqrt{2})/7$.

Answer. We need to construct the side of the square, which is $r\sqrt{\pi}$. Since π itself has become constructible through this convenient (and incorrect) formula, we can construct its square root by the well-known Euclidean construction. We then construct the product by the construction given by Descartes. (See Chapters 10 and 12 for details on how to do these two constructions.

9.2. Prove that the implied Egyptian formula for the volume of a frustum of a square pyramid is correct. If the sides of the upper and lower squares are a and b and the height is h, the implied formula is:

$$V = \frac{h}{3}(a^2 + ab + b^2)\,.$$

Answer. If we imagine the faces of the frustum continued upward to fill in the missing pyramid on top and then slice through the full pyramid diagonally to the two bases, we see that the height H of the complete pyramid satisfies $H : b/\sqrt{2} = (H - h) : a/\sqrt{2}$. That means $H = bh/(b - a)$ and $H - h = ah/(b - a)$. Hence the full pyramid has volume $\frac{1}{3}b^2\bigl(bh/(b-a)\bigr) = \frac{h}{3}\bigl(b^3/(b-a)\bigr)$, and the piece on top has volume $\frac{h}{3}\bigl(a^3/(b-a)\bigr)$. Hence the frustum has volume $\frac{h}{3}\bigl(b^3 - a^3\bigr)/(b - a) = \frac{h}{3}(a^2 + ab + b^2)$.

9.3. Looking at the Egyptian pyramids, with their layers of brick revealed, now that most of the marble facing that was originally present has been removed, one can see that the total number of bricks must be $1 + 4 + 9 + \cdots + n^2$ if the slope (*seked*) is constant. Assuming that the Egyptian engineers had the kind of numerical knowledge that would enable them to find this sum as $\frac{1}{6}n(n + 1)(2n + 1)$, can you conjecture how they may have arrived at the formula for the volume of a frustum? Is it significant that in the only example we have for this computation, the height is 6 units?

Answer. The number of bricks between layer n (from the top) and layer m (from the top), assuming $m > n$ and including layer m but not layer n, is

$$\frac{1}{6}\bigl(2m^3 + 3m^2 + +m - 2n^3 - 3n^2 - n\bigr) = \frac{m-n}{3}\Bigl(m^2 + mn + n^2 + \frac{3}{2}(m + n) + 1\Bigr)\,.$$

If the bricks are very small cubes, say of side t, then the volume of this pyramid is obtained by multiplying this number by t^3. But then $t(m - n) = h$, $t^2m^2 = b^2$, $t^2mn = ab$, and $t^2n^2 = a^2$. We would thus find that

$$V = \frac{h}{3}\Bigl(a^2 + ab + b^2 + \frac{3}{2}(a + b)t + t^2\Bigr)\,.$$

Regarding t as negligibly small, we get the required formula. However, I don't know of any good evidence that the Egyptians did think of the process this way. It looks to me like the kind of retrospective fit that the experienced later generations think up.

9.4. Explain the author's solution of the following problem from the cuneiform tablet BM 85 196. Here the numbers in square brackets were worn off the tablet and have been reconstructed.

> A beam of length 0;30 GAR is leaning against a wall. Its upper end is 0;6 GAR lower than it would be if it were perfectly upright. How far is its lower end from the wall?
>
> Do the following: Square 0;30, obtaining 0;15. Subtracting 0;6 from 0;30 leaves 0;24. Square 0;24, obtaining 0;9,36. Subtract 0;9,36 from [0;15], leaving 0;5,24. What is the square root of 0;5,24? The lower end of the beam is [0;18] from the wall.
>
> When the lower end is 0;18 from the wall, how far has the top slid down? Square 0;18, obtaining 0;5;24. . . .

Answer. The number 0;30 is the hypotenuse of the right triangle formed by the ladder and its projections on the wall and the floor. The 0;24 is the projection on the wall, which was given to be 0;6 GAR less than the full length of the ladder. Hence finding the projection on the floor, which is what the problem asks for, is simply a matter of squaring the hypotenuse, squaring the projection on the wall, subtracting, and then extracting the square root.

9.5. Show that the average of the areas of the two bases of a frustum of a square pyramid is the sum of the squares of the average and semidifference of the sides of the bases. Could this fact have led the Mesopotamian mathematicians astray in their computation of the volume of the frustum? Could the analogy with the area of a trapezoid have been another piece of misleading evidence pointing toward the wrong conclusion?

Answer.

$$\left(\frac{a+b}{2}\right)^2 + \left(\frac{a-b}{2}\right)^2 = \frac{a^2+b^2}{2}\,.$$

It seems very likely to me that, since the area of a trapezoid is the height times the average of the two widths, the Mesopotamians might have believed that the volume of a frustum of a square pyramid is the height times the average of the areas of the two bases. Being the good algebraists that they were, they'd have been pleased to notice that the average is just the sum of the squares of the average and semidifference of the bases.

9.6. The author of the *Zhou Bi Suan Jing* had a numerical method of finding the length of the diagonal of a rectangle of width a and length b, which can be described as follows. Square the sum of width and length, subtract twice the area, then take the square root. Should one conclude from this that the author knew that the square on the hypotenuse was the sum of the squares on the legs?

Answer. This amounts to the assertion that for a right triangle of sides a, b, and c we have

$$c = \sqrt{(a+b)^2 - 2ab}\,.$$

Certainly that formula is *equivalent* to the Pythagorean theorem, but it is not psychologically the same thing. However, one can solve any problem using this formula that one can solve using the Pythagorean theorem, and vice versa. Moreover, I'm sure that if the author of the *Zhou Bi Suan Jing* had seen the Pythagorean theorem stated in the form in which we know it, he would have recognized it immediately as an application of this principle.

9.7. What happens to the estimate of the Sun's altitude (36,000 km) given by Zhao Shuang if the "corrected" figure for shadow lengthening (4 *fen* per 1000 *li*) is used in place of the figure of 1 *fen* per 1000 *li*?

Answer. This would mean that the sun was overhead 15,000 *li* to the south and that the sun is only 20,000 *li* high, that is, about 9,000 km.

9.8. The *gougu* section of the *Jiu Zhang Suanshu* contains the following problem:

> Under a tree 20 feet high and 3 in circumference there grows a vine, which winds seven times the stem of the tree and just reaches its top. How long is the vine?

Solve this problem.

Answer. Imagine the tree cut down and rolled out to unwind the vine. The base of the tree will roll out a line 21 feet long (seven times its circumference). The vine will be the hypotenuse of the right triangle formed by this line and the end position of the tree. Hence its length is $\sqrt{(20)^2 + (21)^2} = 29$ feet.

9.9. Another right-triangle problem from the *Jiu Zhang Suanshu* is the following. "There is a string hanging down from the top of a pole, and the last 3 feet of string are lying flat on the ground. When the string is stretched, it reaches a point 8 feet from the pole. How long is the string?" Solve this problem. You can also, of course, figure out how high the pole is from this information.

Answer. We get the equation $c^2 = 8^2 + (c-3)^2$, so that $6c = 8^2 + 9 = 73$. Hence $c = 12\frac{1}{6}$ feet. (The pole is $9\frac{1}{6}$ feet high.)

9.10. A frequently reprinted problem from the *Jiu Zhang Suanshu* is the "broken bamboo" problem: A bamboo 10 feet high is broken and the top touches the ground at a point 3 feet from the stem. What is the height of the break? Solve this problem, which reappeared several centuries later in the writings of the Hindu mathematician Brahmagupta.

Answer. The equation is $x^2 + 9 = (10-x)^2$, which gives $20x = 91$, so $x = 4.5$ feet.

9.11. The *Jiu Zhang Suanshu* implies that the diameter of a sphere is proportional to the cube root of its volume. Since this fact is equivalent to saying that the volume is proportional to the cube of the diameter, should we infer that the author knew both proportions? More generally, if an author knows (or has proved) "fact A," and fact A is logically equivalent to fact B, is it accurate to say that the author knew or proved fact B? (See also Problem 9.6 above.)

Answer. In general, we shouldn't automatically conclude that people who know one of two logically equivalent propositions also know the second. It all depends on how obvious the equivalence is. In this case, I think it is safe to say that the author would have recognized the equivalence immediately.

9.12. Show that the solution to the quadrilateral problem of Sawaguchi Kazuyuki is $u = 9$, $v = 8$, $w = 5$, $x = 4$, $y = \sqrt{(1213 + 69\sqrt{273})/40}$, $z = 10$. (The approximate value of y is 7.6698551.) From this result, explain how Sawaguchi Kazuyuki must have invented the problem and what the two values 60.8 and 326.2 are approximations for. How does this problem illustrate the claim that these challenge problems were algebraic rather than geometric?

Answer. This is simply a matter of verifying that these lengths do satisfy the condition of being sides and diagonals of a quadrilateral. That is most easily done by representing z as the line joining $(0,0)$ and $(10,0)$ in the st-plane. Then the point where v meets x can be taken as one of the points of intersection of the circles $s^2+t^2 = 16$ and $(s-10)^2+t^2 = 81$, say the point with $t < 0$. Subtracting these equations gives the relation $20s = 35$, so that $s = 1.75$ and so $t = -\sqrt{16-(1.75)^2} = -\sqrt{12.9375}$. Similarly, the point where u meets w can be taken as the intersection of the circle $(s-10)^2 + t^2 = 64$ and $s^2 + t^2 = 25$ at a point where $t > 0$. Subtracting these equations gives $20s = 61$, so that $s = 3.05$ and $t = \sqrt{25 - 3.05^2} = \sqrt{15.6975}$. The length of y is then

$$\sqrt{(1.75-3.05)^2 + (\sqrt{12.9375} + \sqrt{15.6975})^2}\,.$$

This works out to be

$$\sqrt{30.325 + 2\sqrt{203.27390625}} = \sqrt{30.325 + 1.725\sqrt{273}} = \sqrt{(1213 + 69\sqrt{273})/40}\,,$$

which is the number given above. The number 60.8 is thus an approximation to $512 - y^3$, and 326.2 is an approximation to $y^3 - 125$. Notice that the sum is $326.2 + 60.8 = 387 = 512 - 125$.

It is evident that Sawaguchi Kazuyuki first drew two triangles sharing a side of length 10 and the other sides of length 5 and 8 on one side and 4 and 9 on the other side, then simply computed the length of the other diagonal to get y. Then, to make an interesting problem, he gave the data in the form of differences of cubes and challenged others to uncover his tracks.

9.13. How is it possible that some Japanese mathematicians believed the area of the sphere to be one-fourth the square of the circumference, that is, $\pi^2 r^2$ rather than the true value $4\pi r^2$? Smith and Mikami (*1914*, p. 75) suggest a way in which this belief might have appeared plausible. To explain it, we first need to see an example in which the same line of reasoning really does work.

By imagining a circle sliced like a pie into a very large number of very thin pieces, one can imagine it cut open and all the pieces laid out next to one another, as shown in Fig. 17. Because these pieces are very thin, their bases are such short segments of the circle that each base resembles a straight line. Neglecting a very tiny error, we can say that if there are n pieces, the base of each piece is a straight line of length $2\pi r/n$. The segments are then essentially triangles of height r (because of their thinness), and hence area $(1/2)\cdot(2\pi r^2)/n$. Since there are n of them, the total area is πr^2. This heuristic argument gives the correct result. In fact, this very figure appears in a Japanese work from 1698 (Smith and Mikami, *1914*, p. 131).

Now imagine a hemispherical bowl covering the pie. If the slices are extended upward so as to slice the bowl into equally thin segments, and those segments are then straightened out and arranged like the segments of the pie, they also will have bases equal to $\frac{2\pi r}{n}$, but their height will be one-fourth of the circumference, in other words, $\pi r/2$, giving a total area for the hemisphere of $(1/2)\cdot\pi^2 r^2$. Since the area is $2\pi r^2$, this would imply that $\pi = 4$. What is wrong with the argument? How much error would there be in taking $\pi = 4$?

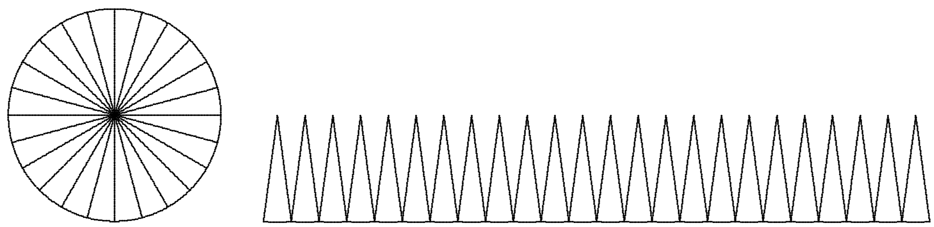


FIGURE 17. A disk cut into sectors and opened up.

Answer. The difficulty with the analogy is that a disk is flat and a sphere is curved. One cannot flatten the sphere without distorting its area. (However, one *can* project it into a cylinder in a way that preserves area.)

9.14. What is the justification for the statement by the historian of mathematics T. Murata that Japanese mathematics (*wasan*) was not a science but an art?

Answer. Mainly, I think, the difference lies in the applications made of the geometry and algebra. An esthetically pleasing geometric figure was esteemed, especially if it led to interesting algebra. The mathematics was not applied to solve any physical or practical problems.

9.15. Show that Aryabhata's list of sine differences can be interpreted in our language as the table whose nth entry is

$$3438\left[\sin\left(\frac{n\pi}{48}\right)-\sin\left(\frac{(n-1)\pi}{48}\right)\right].$$

Use a computer to generate this table for $n = 1, \ldots, 24$, and compare the result with Aryabhata's table.

Answer. The increment in the table entries (225 minutes or 3.75 degrees) is exactly $\pi/48$. When I generated the table with *Mathematica*, I got the numbers 224.856, 223.893, 221.971, 219.099, 215.289, 210.557, 204.923, 198.411, 191.05, 182.871, 173.909, 164.202, 153.792, 142.724, 131.044, 118.803, 106.053, 92.8493, 79.248, 65.3072, 51.0868, 36.6476, 22.0515, 7.36102. Rounding to the nearest integer gives 225, 224, 222, 219, 215, 211, 205, 198, 191, 183, 174, 164, 154, 143, 131, 119, 106, 93, 79, 65, 51, 37, 22, 7. The differences were noted in the text: Aryabhata's 210 should be 211 (but the actual error is only .557) and his 199 should be 198 (again, the actual error is only .589).

9.16. If the recursive procedure described by Aryabhata is followed faithfully (as a computer can do), the result is the following sequence.

225, 224, 222, 219, 215, 210, 204, 197, 189, 181, 172,
162, 151, 140, 128, 115, 102, 88, 74, 60, 45, 30, 15, 0

Compare this list with Aryabhata's list, and note the systematic divergence. These differences should be approximately 225 times the cosine of the appropriate angle. That is, $d_n \approx 225 \cdot \cos\bigl(225(n+0.5)\text{ minutes}\bigr)$. What does that fact suggest about the source of the systematic errors in the recursive procedure described by Aryabhata?

Answer. The differences form the sequence 0, 0, 0, 0, 0, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 4, 4, 5, 5, 5, 6, 7, 7, 7. What we have here is an algebraic function attempting to approximate a transcendental function and losing track of it. If Aryabhata did use this formula, he must have started over with a new "seed" several times in order to avoid the error accumulation shown here.

9.17. Use Aryabhata's procedure to compute the altitude of the Sun above the horizon in London (latitude $51^\circ\,32'$) at 10:00 AM on the vernal equinox. Assume that the sun rises at 6:00 AM on that day and sets at 6:00 PM.

Answer. By the formula the altitude is θ, where

$$\theta = \arcsin\bigl(\sin(51^\circ\,32')\sin(60^\circ)\bigr) = 42^\circ\,41'\,35''\,.$$

(The sun is 60° along its daily arc in the sky at 10:00 AM on the vernal equinox.)

9.18. Why is it necessary that a quadrilateral be inscribed in a circle in order to compute its diagonals knowing the lengths of its sides? Why is it not possible to do so in general?

Answer. The diagonals themselves are not determined by the lengths of the sides, since a quadrilateral frame with pivots at its corners can change its shape. Knowing that all four vertices lie on a circle makes the problem determinate, since the fourth vertex must be at the intersection of the circle circumscribed about the other three vertices and a circle having center at one of those vertices.

9.19. Show that the formula given by Brahmagupta for the area of a quadrilateral is correct if and only if the quadrilateral can be inscribed in a circle.

Answer. For a quadrilateral of sides a, b, c, and d Brahmagupta's formula can be written as the equation

$$16A^2 = 8abcd + 2a^2b^2 + 2a^2c^2 + 2a^2d^2 + 2b^2c^2 + 2b^2d^2 + 2c^2d^2 - a^4 - b^4 - c^4 - d^4.$$

Now a necessary and sufficient condition for a quadrilateral to be inscribed in a circle is that one pair of opposite angles be supplementary. (It then follows that both pairs of opposite angles have this property, since the four angles taken together must sum to four right angles.) Considering a quadrilateral with sides of length a and b on one side of a diagonal of length e and sides of length c and d on the other side, the condition that the angles on opposite sides be supplementary says that their cosines must be negatives of each other. Using the law of cosines, we find

$$a^2 + b^2 - 2ab\cos\theta = e^2 = c^2 + d^2 - 2cd\cos\varphi.$$

Now if θ and φ total two right angles, we must have

$$\cos\theta = -\cos\varphi,$$

and therefore

$$a^2 + b^2 - c^2 - d^2 = 2(ab + cd)\cos\theta,$$

so that

$$\cos\theta = \frac{a^2 + b^2 - c^2 - d^2}{2(ab + cd)}.$$

Now the area of the quadrilateral is

$$A = \frac{1}{2}(ab\sin\theta + cd\sin\varphi).$$

Hence the condition for the vertices to lie on a circle is that

$$A = \frac{1}{2}(ab + cd)\sin\theta =$$

$$\frac{1}{2}(ab + cd)\sqrt{1 - \cos^2\theta} = \sqrt{\left(\frac{ab + cd}{2}\right)^2 - \left[\frac{(a^2 + b^2) - (c^2 + d^2)}{4}\right]^2}.$$

or

$$16A^2 = 4(ab + cd)^2 - [(a^2 + b^2) - (c^2 + d^2)]^2.$$

Expanding the two squares in this last expression and gathering like terms results in precisely the formula of Brahmagupta.

9.20. Imagine a sphere as a polyhedron having a large number of very small faces. Deduce the relation between the volume of a sphere and its area by considering the pyramids obtained by joining the points of each face to the center of the sphere.

Answer. The volume of each pyramid is one-third the area of the base times its height. Now the height of each these pyramids is approximately the radius of the sphere. Hence the volume of each pyramid is the one-third the radius times the area of its base. Summing all of these volumes, we find that the volume of the sphere is one-third the radius time the area of the sphere. That is, $V = \frac{1}{3}Ar$.

CHAPTER 10

Euclidean Geometry

10.1. Show how it would be possible to compute the distance from the center of a square pyramid to the tip of its shadow without entering the pyramid, after first driving a stake into the ground at the point where the shadow tip was located at the moment when vertical poles cast shadows equal to their length.

Answer. One could sight from the stake toward the center (apex) of the pyramid and note where the line of sight intersects the side of the base. It is easy to measure the distance from there to the midpoint of the side of the pyramid, then use the Pythagorean theorem to work out the distance from that point to the center of the pyramid. (One leg is half the side of the pyramid. The other is the distance from the point in question to the midpoint of the side.) Once that distance is found, simply add it to the measured distance from the stake to the point on the side of the pyramid.

10.2. Describe a mechanical device to draw the quadratrix of Hippias. You need a smaller circle of radius $2/\pi$ times the radius that is rotating, so that you can use it to wind up a string attached to the moving line; or conversely, you need the rotating radius to be $\pi/2$ times the radius of the circle pulling the line. How could you get such a pair of circles?

Answer. Wind a string around the circle, then stretch it out to a straight line. You now have a line of length $2\pi r$. Since you want $2r/\pi$, observe that it is $4r^2$ divided by this length. Thus you need to construct a rectangle on the side of length $2\pi r$ whose area equals the square on the diameter of the given circle. That is a classical Euclidean construction (application without defect or excess).

10.3. Prove that the problem of constructing a rectangle of prescribed area on part of a given base a in such a way that the defect is a square is precisely the problem of finding two numbers given their sum and product (the two numbers are the lengths of the sides of the rectangle). Similarly, prove that the problem of application with square excess is precisely the problem of finding two numbers (lengths) given their difference and product.

Answer. Let the two parts of the line have lengths x and y. Then xy is the given area and $x + y$ is the given line. Likewise, if the length of the base of the rectangle in the application with excess problem is x and its excess over the given line is y, then $x - y$ is the given line and xy is the given area.

10.4. Show that the problem of application with square excess has a solution for any given area and any given base. What restrictions are needed on the area and base in order for the problem of application with square defect to have a solution?

Answer. The corner of the rectangle opposite to one end of the given line lies on a ray emanating from the other end of the line and making an angle of $135°$ with the given line. As a point on this ray moves away from the intersection with the given line, the area of

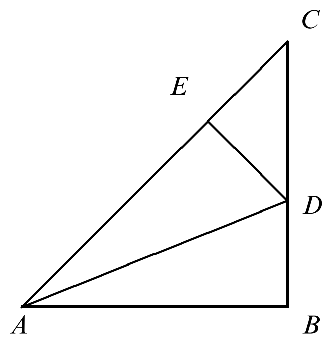


FIGURE 24. Diagonal and side of a square.

the rectangle increases continuously from zero to infinity. Hence there is a unique point at which it equals the given area.

As for application with defect, the angle is 45°, and the area increases as the point moves along the ray only until it is directly above the midpoint of the given line. Then it decreases to zero when the point is above the opposite endpoint. Hence the maximum area that can be constructed is the square on half the line. In other words, the pair of equations $x + y = a$, $xy = A$ has positive solutions x, y if and only if $A \leq \frac{1}{4}a^2$ That, of course is well, known, since the equations imply $x(a - x) = A$, that is, $x^2 - ax + A = 0$, and the discriminant condition is $a^2 - 4A \geq 0$.

10.5. Use an argument similar to the argument in Chapter 8 showing that the side and diagonal of a pentagon are incommensurable to show that the side and diagonal of a square are incommensurable. That is, show that the Euclidean algorithm, when applied to the diagonal and side of a square, requires only two steps to produce the side and diagonal of a smaller square, and hence can never produce an equal pair. To do so, refer to Fig. 24.

In this figure $AB = BC$, angle ABC is a right angle, AD is the bisector of angle CAB, and DE is drawn perpendicular to AC. Prove that $BD = DE$, $DE = EC$, and $AB = AE$. Then show that the Euclidean algorithm starting with the pair (AC, AB) leads first to the pair $(AB, EC) = (BC, BD)$, and then to the pair $(CD, BD) = (CD, DE)$, and these last two are the diagonal and side of a square.

Answer. Side AD is common to triangles ADE and ADB. Since AD is the bisector of $\angle CAB$, $BD \perp AB$, and $DE \perp AC$, it follows that triangles ADE and ADB are congruent by angle–angle–side. Since $\angle CDE$ must equal $\angle CAB$ (because both must be complementary to $\angle C$ in their respective right triangles), it follows that $\angle CDE$ is congruent to $\angle C$, hence that triangle DCE is isosceles. Therefore $EC = ED = BD$. We have $AB = AE$ by the congruence of triangles ADB and ADE.

Starting the Euclidean algorithm with the pair (AC, AB), we get $(AC - AB, AB) = (AC - AE, AB) = (EC, AB) = (BD, BC)$. Since $BC > BD$, our next pair is $(BC - BD, BD) = (CD, BD) = (CD, DE)$, which, as asserted, form the diagonal and side of a square.

10.6. It was stated above that Thales might have used the Pythagorean theorem in order to calculate the distance from the center of the Great Pyramid to the tip of its shadow. How could this distance be computed without the Pythagorean theorem?

Answer. Draw as much as you can of a right triangle having the line from the center to the tip of the shadow as its hypotenuse and one leg parallel to a side of the pyramid. You will be able to draw all of this side. (The other leg will of course be perpendicular to the side, and you won't be able to draw all of it.) Reflect this triangle about the point where the tip of the shadow is. You can draw as long an extension of the hypotenuse as you like. If you reflect all of the side parallel to the side of the pyramid, then draw a perpendicular to that line at the endpoint of the extension, the point where the perpendicular intersects the extension of the hypotenuse will determine the third vertex of a triangle congruent to the one you can't measure.

10.7. State the paradoxes of Zeno in your own words and tell how you would have advised the Pythagoreans to modify their system in order to avoid these paradoxes.

Answer. My response to this question is already contained in the text, since the paradoxes have been stated there in my words. In all cases, I think, the disconnect between verbal reasoning and geometric intuition arises because the concept of a continuum requires a set that is *uncountable* and an uncountable set cannot be "approximated" by a large finite set. It is a qualitatively different thing. The qualitative difference shows up in the notion of geometric dimension and the consequent differences among length, area, and volume. These different dimensions lose the Archimedean property: No multiple of a point can ever exceed a curve, surface, or solid. No multiple of a curve can ever exceed a surface or solid, and no multiple of a surface can ever exceed a solid. Hence one cannot think of a line as being a bunch of points stacked up one on top of another. The intuitive attempt to do that has to be replaced by a limiting process that works "down from above," taking lines or curves of finite length and letting that length go to zero. That way, the object you are using to reach the infinitesimal level is comparable to the object you hope to synthesize from those infinitesimals.

10.8. Do we share any of the Pythagorean mysticism about geometric shapes that Proclus mentioned? Think of the way in which we refer to an honorable person as *upright*, or speak of getting a *square deal*, while a person who cheats is said to be *crooked*. Are there other geometric images in our speech that have ethical connotations?

Answer. A person who speaks truthfully and frankly is said to be a *straight-talker*. Uncomplicated people who are regarded as dull are sometimes said to be *one-dimensional*. A broadly educated person is said to be *well-rounded*. Words such as *rectitude*, coming from the Latin, have similar roots. Topological notions enter ethics in such words as *integrity* (wholeness) and *duplicity*, the image being that a person of integrity is consistent and can be relied upon, while a duplicitous (two-faced) person may appear to be of one opinion in one context and an opposite opinion in some other. An insensitive person is sometimes said to be *obtuse* (as opposed to *acute*).

10.9. In the Pythagorean tradition there were two kinds of mathematical activity. One kind, represented by the attempt to extend the theory of the transformation of polygons to circles and solid figures, is an attempt to discover new facts and enlarge the sphere of mathematics—to generalize. The other, represented by the discovery of incommensurables, is an attempt to bring into sharper focus the theorems already proved and to test the underlying assumptions of a theory—to rigorize. Are these kinds of activity complementary, opposed, or simply unrelated to each other?

Answer. They are partly independent, partly interactive. Some mathematicians gravitate toward the philosophical end of the subject, others are caught up in its intuitive aspects and like to create new theories. Sometimes new theories are created in order to take account of criticism, and sometimes reflection on the inner workings of a subject lead to useful generalizations of it, as the basic underlying principles are isolated.

10.10. Hippocrates' quadrature of a lune used the fact that the areas of circles are proportional to the squares on their radii. Could Hippocrates have known this fact? Could he have proved it?

Answer. That depends on what you mean by knowing and what you mean by proof. Hippocrates was a late Pythagorean, who probably lived in the second half of the fifth century BCE. He would have inherited the semi-formal approach to geometry that was partly intuitive, partly logical. The more intense logical rigor that came after the criticism of Zeno and the work of Plato's students wouldn't have been part of it. In particular, he lived long before Eudoxus gave a theory of proportion that could handle incommensurable magnitudes.

10.11. Plato apparently refers to the famous 3–4–5 right triangle in the *Republic*, 546c. Proclus alludes to this passage in a discussion of right triangles with commensurable sides. We can formulate the recipes that Proclus attributes to Pythagoras and Plato respectively as

$$(2n+1)^2 + (2n^2+2n)^2 = (2n^2+2n+1)^2$$

and

$$(2n)^2 + (n^2-1)^2 = (n^2+1)^2.$$

Considering that Euclid's treatise is regarded as a compendium of Pythagorean mathematics, why is this topic not discussed? In which book of the *Elements* would it belong?

Answer. As we saw in the text, Euclid also does not discuss figurate numbers, another topic of interest to the Pythagoreans. Euclid was much more interested in the theory of proportion than anything else. Once you have that theory and the general Pythagorean theorem, integer-sided triangles are a topic of limited interest. If he had discussed this topic, it would have been in Books 7–9.

10.12. Proposition 14 of Book 2 of Euclid shows how to construct a square equal in area to a rectangle. Since this construction is logically equivalent to constructing the mean proportional between two line segments, why does Euclid wait until Book 6, Proposition 13 to give the construction of the mean proportional?

Answer. The equivalence of the two things is not obvious and cannot be proved without the Eudoxan theory of proportion. Euclid does prove the equivalence in Book 6, Proposition 16.

10.13. Show that the problem of squaring the circle is equivalent to the problem of squaring one segment of a circle when the central angle subtended by the segment is known. (Knowing a central angle means having two line segments whose ratio is the same as the ratio of the angle to a full revolution.)

Answer. It is very easy, and a direct application of the Eudoxan definition to show that the area of a *sector* of a circle is directly proportional to the angle it subtends. Now suppose that a and b are lines whose ratio $a : b$ equals the ratio of the whole circle to the sector and that s is the side of a square equal to the area of the sector. We can find t, the side of a square equal to the whole circle, as follows.

Let m be the mean proportional between a and b, and construct on side m a rectangle equal to the rectangle whose sides are a and s. Then t is the other side of the rectangle just constructed. For $tm = as$, so that $t : s = a : m$. But the duplicate ratio $(a : m).(a : m)$ is $a : b$, since $m : b = a : m$. It follows that the duplicate ratio $(t : s).(t : s)$ is also $a : b$, and hence the squares on sides t and s are in the ratio $a : b$. Since the square on s equals the sector, it follows that the square on t equals the whole circle. The same argument shows that one can square the sector if one knows how to square the circle.

Now, if we can square the sector subtended by the segment, then the segment, being the difference between the sector and a triangle, can also be squared. The procedure is as follows. Construct a square equal to the sector. Draw a semicircle having one side of that square as diameter. Then, from one end of the diameter, mark off a chord equal to the side of the square whose area is that of the triangular part of the sector. Connect the other end of the chord to the other end of the diameter. That second chord is the side of a square equal to the segment. Conversely, if we can square the segment, since we can also square the triangle that complements the segment to make a sector, we can put the two squares corner to corner so that their sides form the legs of a right triangle, and the square on the hypotenuse will then be equal to the sector.

10.14. Referring to Fig. 18, show that all the right triangles in the figure formed by connecting B' with C, C' with K, and K' with L are similar. Write down a string of equal ratios (of their legs). Then add all the numerators and denominators to deduce the equation

$$(BB' + CC' + \cdots + KK' + LM) : AM = A'B : BA.$$

Answer. We have duplicated Fig. 18 here and added the letters O, P, Q, R, S, and T to make it easier to refer to the triangles we need. Each of these right triangles has an acute angle inscribed in an arc equal to the arc $\overset{frown}{AB}$. Thus all the triangles AOB, POB', PQC, RQC', RSK, TSK' and TML are similar to one another and to ABA'. Thus we have

$$\frac{OB}{AO} = \frac{OB'}{OP} = \frac{QC}{PQ} = \frac{QC'}{QR} = \frac{SK}{RS} = \frac{SK'}{ST} = \frac{LM}{TM}\,.$$

Therefore, adding numerators to numerators and denominators to denominators, we find that

$$\frac{OB + OB' + QC + QC' + SK + SK' + LM}{AM} = \frac{A'B}{AB}\,.$$

This is the result asked for, since $OB + OB' = BB'$, and so on. Actually, for purposes of completing Archimedes' argument, it makes more sense to leave the expression as it is here. For, according to Archimedes' lemma the frustum of a cone generated by revolving, say $B'C'$ about the diameter AA' has area equal to a circle whose radius is the mean proportional between BC and the sum $OB' + QC$. Since the area of a circle is proportional (with constant π) to the square of the radius, that area is just π times $BC(OB' + QC)$. Since $AB = BC = CK = KL$, cross-multiplying in this last equation and replacing AB successively by BC, CK, and KL, we find that the area generated by revolving $ABCKL$ about AA' is π times

$$AB \cdot OB + BC(OB' + QC) + CK(QC' + SK) + KL(SK' + LM) = AM \cdot A'B\,.$$

But if the side AB of the polygon is taken sufficiently small, both AM and $A'B$ will be as close to the diameter as we like, and the area of the revolved polygon will be as close to the area of the sphere as we like. Hence the area of the sphere is π times the square of its diameter.

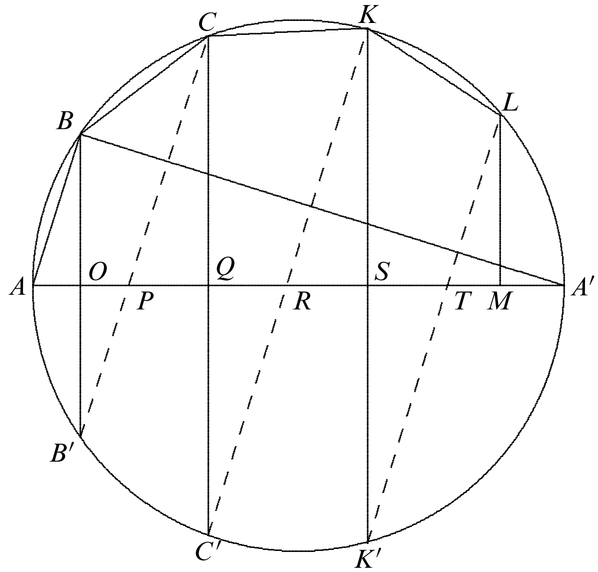
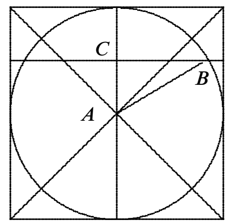


FIGURE 18. Finding the surface area of a sphere.

Archimedes would have said that if the square on the radius of a given circle equals the sum of the squares on the radii of a (finite) set of circles, then the area of the given circle equals the sum of the areas of the set of circles. This equation is thus interpreted by saying that the square of the radius of a circle equal to the area obtained by revolving the polygon about AA' is $AM \cdot A'B$, and (in the limit, as we would say) the area of the sphere equals the area of a circle whose squared radius equals the squared diameter of the sphere, that is, a circle whose radius equals the diameter of the sphere. But of course, as Archimedes knew and stated, the area of such a circle is four times the area of a great circle on the sphere.

10.15. Show that Archimedes' result on the relative volumes of the sphere, cylinder, and cone can be obtained by considering the cylinder, sphere and double-napped cone formed by revolving a circle inscribed in a square about a midline of the square, the cone being generated by the diagonals of the square. In this case the area of a circular section of the cone plus the area of the same section of the sphere equals the area of the section of the cylinder since the three radii form the sides of a right triangle. The radius of a section of the sphere cuts off a segment of the axis of rotation from the center equal to the radius of the section of the cone, since the vertex angle of the cone is a right angle. These two segments form the legs of a right triangle whose hypotenuse is a radius of the sphere, which is equal to the radius of the section of the cylinder.

Answer. In the vertical cross section of such a solid (see the accompanying figure) the square is the section of the cylinder, the circle the section of the sphere, and the crossed diagonals the section of the double-napped cone. The horizontal line one-fourth of the distance from the top of the cylinder represents a horizontal cross section, whose intersections with the three solids give three concentric circles of different radii. Since the areas of these circles are proportional to the squares on the radii, we have only to observe that the square on the radius of the cylindrical section equals the sum of the squares on the radii of the sections of the sphere and cone. (Draw the radius AB from the center of the figure to the



Sections of sphere, cylinder, and cone

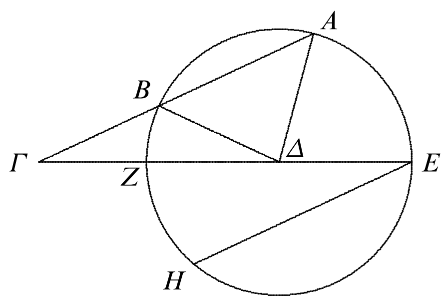


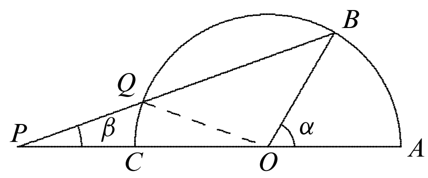
FIGURE 25. Archimedes' trisection of an angle: $\angle A\Gamma\Delta = \frac{1}{3}\angle A\Delta E$.

point where the horizontal line meets the circle. Together with the horizontal arm BC and the vertical arm AC it forms a right triangle ABC. The vertical arm AC equals the radius of the conical cross section, since the generating lines of the cone form a 45°-degree angle with the axis of the cone. The horizontal arm BC is the radius of the cross-section of the sphere, and the hypotenuse AB is the radius of the cylinder.

The conclusion is that the volume of the cylinder is the sum of the volumes of the cone and the sphere.

10.16. A minor work attributed to Archimedes called the *Book of Lemmas* contains an angle trisection. In Fig. 25 we are given an acute angle $\angle A\Delta E$, whose trisection is required. We draw a circle of any radius r about Δ, the vertex of the angle. Then, using a straightedge, we mark off on it two points P and Q separated by the distance r. Setting the straightedge down so that P is at point Γ on the extension of the diameter $E\Delta Z$, Q is at point B on the circle, and the point A is also on the edge of the straightedge, we draw the line $A\Gamma$. By drawing EH parallel to $A\Gamma$, we get $\angle A\Gamma E = \angle \Gamma EH$. By joining ΔB, we obtain the isosceles triangle $\Gamma B\Delta$. Now since $\angle B\Delta Z$ is a central angle on the arc $\overset{\frown}{BZ}$ and is equal to $\angle B\Gamma\Delta$, which is equal to $\angle ZEH$, which is inscribed in the arc $\overset{\frown}{ZH}$, it follows that $\overset{\frown}{ZH} = 2\ \overset{\frown}{BZ}$. Since the arcs $\overset{\frown}{AE}$ and $\overset{\frown}{BH}$ are equal (being cut off by parallel chords), we now get $\overset{\frown}{AE} = \overset{\frown}{BH} = 3\ \overset{\frown}{BZ}$. Therefore, $\angle A\Gamma E = \angle B\Delta Z = \frac{1}{3}\angle A\Delta E$.

Why is this construction *not* a straightedge-and-compass trisection of the angle, which is known to be impossible? How does it compare with the *neûsis* trisection shown above?



Show how to obtain this same result more simply by erasing everything in the figure below the diameter of the circle.

Answer. Letters here refer to the accompanying figure. We are given an acute angle $\alpha = \angle AOB$, whose trisection is required. We draw a circle ABC of any radius r about O, the vertex of the angle. Then, using a straightedge, we mark off on it two points P and Q separated by the distance r. Setting the straightedge down so that P is on the extension of the diameter CA, Q is on the semicircle ABC, and the point B is also on the edge of the straightedge, we draw the line PB, which contains the point Q. By drawing the radius QO we obtain two isosceles triangles OQP and QOB. The equal angles of the first of these will be denoted β, and since the exterior angle of a triangle equals the sum of the two opposite interior angles, it follows that the equal angles of the second are equal to 2β. Therefore $\angle BPO = \beta$, $\angle PBO = 2\beta$, and again by the exterior angle theorem $\alpha = 3\beta$. That is, we have constructed an angle β equal to one-third of α.

The construction fails because the line PB is not determined, except visually. In geometry a line can be determined in only three ways: 1) by knowing two points on it; 2) by knowing one point on it and the angle it makes with a second line through that point; 3) by knowing one point on it and a line that is parallel to it. None of these conditions is met in the present case.

10.17. Show that the problem of increasing the size of a sphere by half is equivalent to the problem of two mean proportionals (doubling the cube).

Answer. Given a sphere of radius r, one needs to find the radius s of a sphere that is half again as large. Since spheres are in proportion to the triplicate ratio of their radii, it would suffice to find two lines s and t such that $r : s = s : t = t : \frac{3}{2}r$. Conversely, if one could find the radius s, one could construct the third proportional t such that $r : s = s : t$ (Euclid, Book 6, Proposition 11), and then the third proportional u such that $s : t = t : u$. But then $r : u$ would be the triplicate ratio of $r : s$ and that triplicate ratio is $2 : 3$ by the assumption on s. Hence $u = \frac{3}{2}r$.

10.18. A circle can be regarded as a special case of an ellipse. What is the *latus rectum* of a circle?

Answer. The *latus rectum* of a circle is its diameter, since the perpendicular from a circle to its diameter is the mean proportional between the segments of the diameter.

10.19. When the equation $y^2 = Cx - kx^2$ is converted to the standard form

$$\frac{(x-h)^2}{a^2} + \frac{y^2}{b^2} = 1,$$

what are the quantities h, a, and b in terms of C and k?

Answer. This is simple algebra: $a = \frac{C}{2k}$, $b = \frac{C}{2\sqrt{k}}$.

10.20. Show from Apollonius' definition of the foci that the product of the distances from each focus to the ends of the major axis of an ellipse equals the square on half of the minor axis.

Answer. In general, any rectangle deficient by a square has area equal to the product of the two line segments into which it divides the line segment to which it is applied. Apollonius is asserting that (what we call) the foci are points at which this product is one-fourth the product of the major axis and the latus rectum. By the geometric way in which the ellipse is defined, it is clear that the square of the ordinate at the midpoint of the major axis (that is, the square on half of the minor axis) will be exactly one-fourth of the rectangle on the major axis and the latus rectum. (Simply put, if l is the latus rectum, and a and b are half of the major and minor axes respectively, then $2b^2 = al$. Note that for a circle, where a and b are both equal to the radius, this formula gives l as the diameter of the circle.) Hence the assertion follows.

10.21. We have seen that the three- and four-line locus problems have conic sections as their solutions. State and solve the two-line locus problem. You may use modern analytic geometry and assume that the two lines are the x axis and the line $y = ax$. The locus is the set of points whose distances to these two lines have a given ratio. What curve is this?

Answer. The distance from a general point (x, y) to the line $ax + by = c$ is well-known to be $|ax + by - c|/\sqrt{a^2 + b^2}$. Hence the general equation is

$$|y| = \frac{r}{\sqrt{a^2 + 1}}|ax - y| = q|ax - y|\,,$$

where r is the ratio of the two distances and $q = r/\sqrt{a^2 + 1}$. By squaring both sides, transposing the left-hand side to the right, and then factoring the difference of the two squares, we obtain the equation

$$[aqx + (1 - q)y][aqx - (1 + q)y] = 0,$$

which represents a pair of lines through the intersection of the two given lines. A pair of intersecting lines is considered a degenerate hyperbola.

10.22. Show that the apparent generality of Apollonius' statement of the three-line locus problem, in which arbitrary angles can be prescribed at which lines are drawn from the locus to the fixed lines, is illusory. (To do this, show that the ratio of a line from a point P to line l making a fixed angle θ with the line l bears a constant ratio to the line segment from P perpendicular to l. Hence if the problem is solved for all ratios in the special case when lines are drawn from the locus perpendicular to the given lines, then it is solved for all ratios in any case.)

Answer. The constant ratio is in fact the sine of the given angle at which the oblique lines are drawn. To avoid using trigonometry, one need only observe that the triangles formed by the fixed line, the perpendicular to it, and the oblique to it, are all right triangles having the given angle as an acute angle. Hence they are all similar, and, in particular the ratio of the two lines in question is the same for all of them.

10.23. Show that the line segment from a point $P = (x, y)$ to a line $ax + by = c$ making angle θ with the line has length

$$\frac{|ax + by - c|}{\sqrt{a^2 + b^2}\sin\theta}.$$

Use this expression and three given lines $l_i : a_i x + b_i y = c_i$, $i = 1, 2, 3$, to formulate the three-line locus problem analytically as a quadratic equation in two variables by setting the square of the distance from (x, y) to line l_1 equal to a constant multiple of the product of the distances to l_2 and l_3. Show that the locus passes through the intersection of the line l_1 with l_2 and l_3, but not through the intersection of l_2 with l_3. Also show that its tangent line where it intersects l_i is l_i itself, $i = 2, 3$.

Answer. The first assertion is more or less covered in the preceding two problems. It is standard analytic geometry. The proposed equation has the form

$$(a_1 x + b_1 y - c_1)^2 = r(a_2 x + b_2 y - c_2)(a_3 x + b_3 y - c_3),$$

This is a quadratic expression in x and y, hence represents a conic section. Note that this equation is satisfied when $a_i x + b_i y - c_i = 0$ for $i = 1, 2$ and $i = 1, 3$, but not for $i = 2, 3$ (except in the degenerate case when all three lines are concurrent).

Using implicit differentiation, we find that when $a_i x + b_i y - c_i = 0$ for $i = 1, 2$ but not for $i = 3$, the slope of the tangent line is

$$\frac{dy}{dx} = -\frac{a_2}{b_2},$$

which is exactly the slope of l_2.

10.24. One reason for doubting Cavalieri's principle is that it breaks down in one dimension. Consider, for instance, that every section of a right triangle parallel to one of its legs meets the other leg and the hypotenuse in congruent figures (a single point in each case). Yet the other leg and the hypotenuse are obviously of different lengths. Is there a way of redefining "sections" for one-dimensional figures so that Cavalieri's principle can be retained? If you could do this, would your confidence in the validity of the principle be restored?

Answer. One could define the "zero-dimensional volume" of the point of intersection of two lines as the cosecant of their angle of intersection, so that two lines intersecting at a right angle would have an intersection of zero-dimensional volume 1 and two lines that coincide would have an intersection whose zero-dimensional volume is infinite, as one would expect. Note that the cosecant is the same for any of the four angles formed by two intersecting lines, so that this concept is unambiguously defined. For two intersecting curves one could define the volume to be the volume of the intersection of their tangents at the point of intersection.

This definition would then give consistent results for lines and curves in a plane. Incidentally, it provides a theorem about plane curves: *Let $y = f(x)$ and $y = g(x)$ be plane curves having continuously turning nonvertical tangents at each point $x = c$ for all $c \in [a, b]$ If for all $c \in [a, b]$ the cosecant of the angle of intersection of the curve $y = f(x)$ with the vertical line $x = c$ bears the ratio r to the cosecant of the angle of intersection of the curve $y = g(x)$ with the same line, then the length of the former is r times the length of the latter.* The proof is the observation that the cosecant of the angle in question is the secant of the angle between the tangent line and the horizontal, that is, it is $\sqrt{1 + (f'(c))^2}$ and $\sqrt{1 + (g'(c))^2}$ for the two curves, and the integrals of these two functions give the arc lengths of the two curves.

The need to consider this case points to a perhaps unnoticed assumption in the original statement of the principle and a possibility of generalizing it. The unnoticed assumption was that the sections of the given figures are taken inside a space of the same dimension as the figures themselves. The possibility of considering, for example, one-dimensional sections of two-dimensional figures in three-dimensional space requires some convention such as the one just introduced for zero-dimensional sections of one-dimensional figures in two-dimensional space.

10.25. We know that interest in conic sections *arose* because of their application to the problem of two mean proportionals (doubling the cube). Why do you think interest in them was *sustained* to the extent that caused Euclid, Aristaeus, and Apollonius to write treatises developing their properties in such detail?

Answer. As I hope the discussion above has shown, these objects have properties that are fascinating to contemplate. They are of great esthetic beauty, and in addition suggest applications to many other areas, especially in physics. The discussion in Chapter 12 shows that conic sections were a source of a huge amount of speculation because of their projective properties. That aspect of the subject made them an essential part of geometry right down to the present day.

10.26. Pappus' history of the conics implies that people knew that the ellipse, for example, could be obtained by cutting a right-angled cone with a plane. Can *every* ellipse be obtained by cutting a right-angled cone with a plane? Prove that it can, by showing that any a and b whatsoever in Eq. (2) can be obtained as the section of the right-angled cone whose equation is $y^2 = zx$ by the plane $x = 2a - (a^2z/b^2)$. Then show that by taking $a = eu/(1-e^2)$, $b = a\sqrt{1-e^2}$, $x = w$, $y = v$, where $e = h/w$, you get Eq. (1). [*Hint*: Recall that e is constant in a given conic section. Also, observe that $0 < e < 1$ for a section of an acute-angled cone, since $h = w\tan(\theta/2)$, where θ is the vertex angle of the cone.]

Answer. This is mostly just algebra. Substituting $z = 2b^2/a-(b^2x/a^2)$ yields the equation $y^2 = (2b^2/a)x - b^2x^2/a^2$, which we write as

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

which is

$$\frac{(x-a)^2}{a^2} + \frac{y^2}{b^2} = 1\,.$$

The rest of the problem is very simple algebra.

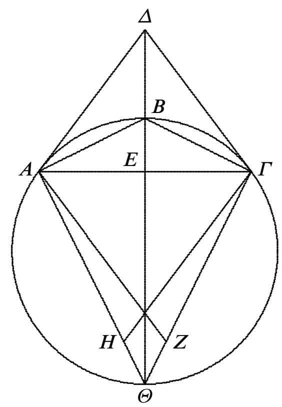
10.27. As we have seen, Apollonius was aware of the string property of ellipses, yet he did not mention that this property could be used to draw an ellipse. Do you think that he did not *notice* this fact, or did he omit to mention it because he considered it unimportant?

Answer. My guess is that he "had bigger fish to fry" at the time. The fact may have been well-known and considered to be non-mathematical, since it involved physical objects and instruments. As such, it would have been out of place in a treatise on pure mathematics. Also, it was not in the main line of his development of the subject, which was to imitate Euclid and find the proportions among the important lines in conic sections.

10.28. Prove Proposition 54 of Book 3 of Apollonius' *Conics* in the special case in which the conic is a circle and the point Θ is at the opposite end of the diameter from B (Fig. 22).

Answer. In this special case, where everything is symmetric, the relation to be proved is

$$\frac{AZ}{A\Gamma} = \frac{A\Delta}{AE}\frac{EB}{B\Delta}\,.$$



Because of the symmetry, $\angle \Delta B = \angle BAE$. (The left-hand side equals half of the central angle on $\widehat{AB}$, and the right-hand side is half of the central angle on $\widehat{B\Gamma}$.) Let us call this angle σ. Then, because $A\Delta$ is parallel to ΓH, it follows that $\angle A\Gamma H$ equals 2σ. But then the bisector of $\angle A\Gamma H$ makes the same angle with $A\Gamma$ that AB makes with $A\Gamma$, namely σ. Hence the bisector of $\angle A\Gamma H$ is parallel to AB. Since AB is perpendicular to $A\Theta$ (the angle ΘAB is inscribed in a semicircle), it follows that the bisector of $\angle A\Gamma H$ is perpendicular to $A\Theta$. Therefore Γ is the vertex angle of the isosceles triangle $A\Gamma H$. But that means $A\Gamma = \Gamma H = AZ$, and hence the relation we need to prove is

$$\frac{AE}{A\Delta} = \frac{EB}{B\Delta}\,.$$

This relation is most easily derived from the law of sines:

$$\frac{EB}{B\Delta} = \frac{EB}{AB}\frac{AB}{B\Delta} = \sin\sigma \cdot \frac{\sin\angle A\Delta\Theta}{\sin\sigma} = \sin\angle A\Delta\Theta = \frac{AE}{A\Delta}\,.$$

CHAPTER 11

Post-Euclidean Geometry

11.1. The figure used by Zenodorus at the main step in his proof of the isoperimetric inequality had been used earlier by Euclid to show that the apparent size of objects is not inversely proportional to their distance. Prove this result by referring to the diagram on the left in Fig. 14. Show that $BE : E\Delta :: AB : Z\Delta :: \Gamma\Delta : Z\Delta$ and that this last ratio is larger than $\overset{\frown}{H\Theta}:\overset{\frown}{Z\Theta}$.

Answer. The first proportion is merely the fact that triangle EBA is similar to triangle $A\Delta Z$. The second is the fact that $AB = \Gamma Z$. As for the inequality, since triangles $EZ\Gamma$ and $E\Delta Z$ have the same altitude (namely $E\Delta$), their areas are proportional to their bases. Thus $\Gamma Z : \Delta Z = \triangle EZ\Gamma : \triangle E\Delta Z >$ Sector $EZH : \triangle E\Delta Z >$ Sector EZH : Sector $A\Theta Z = \overset{\frown}{HZ}:\overset{\frown}{Z\Theta}$. Adding 1 to both sides of the proportion, we get $\Gamma\Delta > Z\Delta > \overset{\frown}{H\Theta}:\overset{\frown}{Z\Theta}$.

11.2. Use the diagram on the right in Fig. 14 to show that the ratio of a larger chord to a smaller is less than the ratio of the arcs they subtend, that is, show that $B\Gamma : AB$ is less than $\overset{\frown}{B\Gamma}:\overset{\frown}{AB}$, where $A\Gamma$ and ΔZ are perpendicular to each other. (Hint: $B\Delta$ bisects angle $AB\Gamma$.) Ptolemy said, paradoxically, that the chord of 1° had been proved "both larger and smaller than the same number" so that it must be *approximately* $1; 2, 50$. Carry out the analysis carefully and get accurate upper and lower bounds for the chord of 1°. Convert

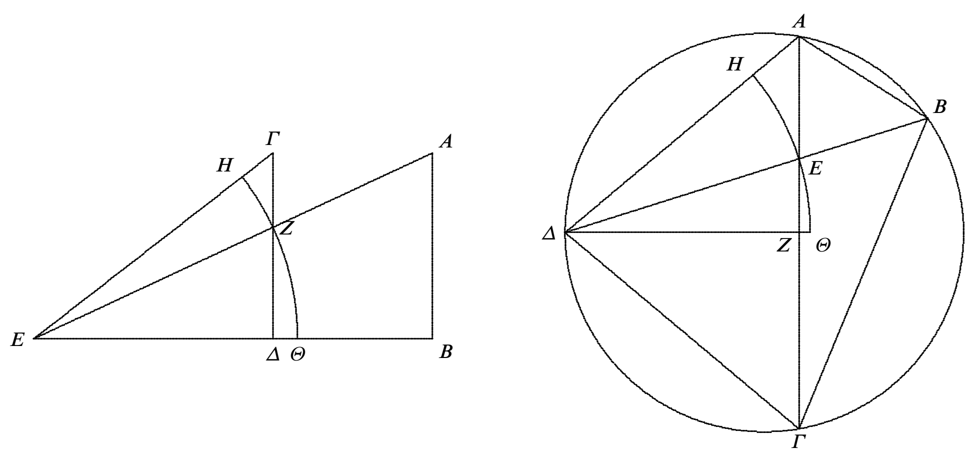


FIGURE 14. Greek use of a fundamental inequality. Left: from Euclid's *Optics*. Right: from Ptolemy's *Almagest*.

this result to decimal notation, and compare with the actual chord of 1° which you can find from a calculator. (It is $120 \sin\left(\frac{1}{2}^\circ\right)$.)

Answer. This question is poorly worded, the result of inattention in proofreading. It had originally been in the text itself. When I moved it to the exercises, I should have stated as part of the construction that $B\Delta$ is the bisector of $\angle AB\Gamma$ and that ΔZ is drawn perpendicular to $A\Gamma$. In any case, we can see that the figure on the left is repeated on the right, and so that argument shows that $AE : AZ > \overset{\frown}{HE} : \overset{\frown}{H\Theta}$. In terms of angles, this says $AE : \frac{1}{2}A\Gamma > \angle A\Delta B : \angle A\Delta Z$. Because of the bisection, $\overset{\frown}{A\Delta} = \overset{\frown}{\Gamma\Delta}$, and therefore Δ is the midpoint of $\overset{\frown}{A\Gamma}$. It then follows that Z is the midpoint of $A\Gamma$. But then, ΔZ extended is a diameter of the circle, and hence $\angle A\Delta Z$ is half of $\angle A\Delta\Gamma$. Thus,

$$AE : \frac{1}{2}A\Gamma > \angle A\Delta B : \frac{1}{2}\angle A\Delta\Gamma\,.$$

Hence we can clear out the fractions and write this as

$$A\Gamma : AE < \angle A\Delta\Gamma : \angle A\Delta B\,.$$

Subtracting 1 from both sides of this inequality gives

$$E\Gamma : AE < \angle E\Delta\Gamma : \angle A\Delta B\,.$$

But, since $B\Delta$ is the bisector of angle $AB\Gamma$, it follows that $AB : AE = \Gamma B : E\Gamma$.[1] And since inscribed angles are proportional to their arcs, we get

$$B\Gamma : AB < \overset{\frown}{B\Gamma} : \overset{\frown}{AB}\,,$$

which was to be proved.

Now Ptolemy had said that the chord of $\frac{3}{4}^\circ$ was $0;\,47,\,8$ where the radius is 60 and that the chord of $1\frac{1}{2}^\circ$ was $1;\,34,\,15$. Hence it followed that the chord $\mathrm{Ch}\,(1^\circ)$ had to satisfy

$$1;\,2,\,50 = \frac{2}{3}(1;\,34,\,15) < \mathrm{Ch}\,(1^\circ) < \frac{4}{3}(0;\,47\,8) = (1;\,20)(0;\,47\,8) = 1;\,2,\,50,\,40\,.$$

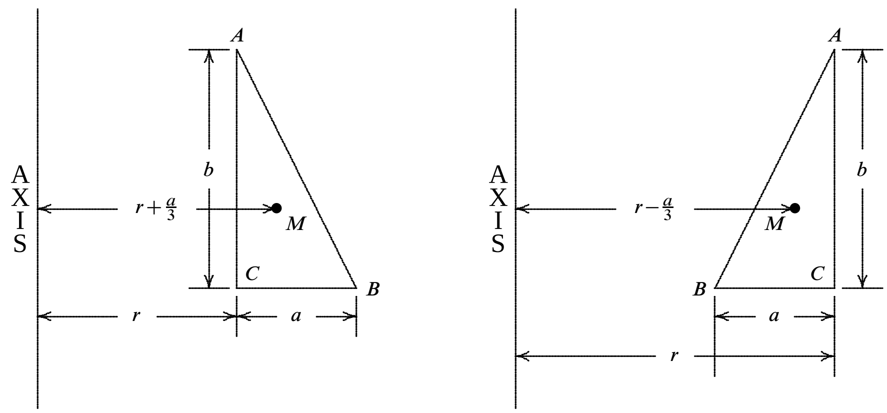
Actually, Ptolemy didn't have it quite right. The chord of $1\frac{1}{2}^\circ$ is about $1;\,34,\,14,\,42,\,19$ and the chord of $\frac{3}{4}^\circ$ is about $0;\,47,\,7,\,24,\,47$. Thus the left-hand estimate here is not quite proved by Ptolemy's argument, and in fact it is not correct. The chord of 1° is really about $1;\,2,\,49,\,51,\,48$, which is ever so slightly smaller than Ptolemy claimed it was. But let us remember that he *did* only claim that this was an approximate value, and that is true.

11.3. Let A, B, C, and D be squares such that $A : B :: C : D$, and let r, s, t, and u be their respective sides. Show that $r : s :: t : u$ by strict Eudoxan reasoning, giving the reason for each of the following implications. Let m and n be any positive integers. Then

$$mr > ns \Rightarrow m^2A > n^2B \Rightarrow m^2C > n^2D \Rightarrow mt > nu\,.$$

Answer. The first implication is simply a matter of squaring the inequality. The second is the definition of the proportion $A : B :: C : D$, and the third is a matter of taking the square root of the preceding inequality. Taken altogether these inequalities (for all m and n constitute a proof that $r : s :: t : u$.

[1] This is Proposition 3 of Book 6 of Euclid, but is not hard to prove, since the triangles $\Delta E\Gamma$ and AEB are similar, as are $\Delta\Gamma B$ and $\Delta E\Gamma$. Hence $\Delta\Gamma : \Delta E = AB : AE$ and $\Delta\Gamma : \Delta E = \Gamma B : E\Gamma$.



Pappus' theorem

11.4. Sketch a proof of Pappus' theorem on solids of revolution by beginning with right triangles having a leg parallel to the axis of rotation, then progressing to unions of areas for which the theorem holds, and finally to general areas that can be approximated by unions of triangles.

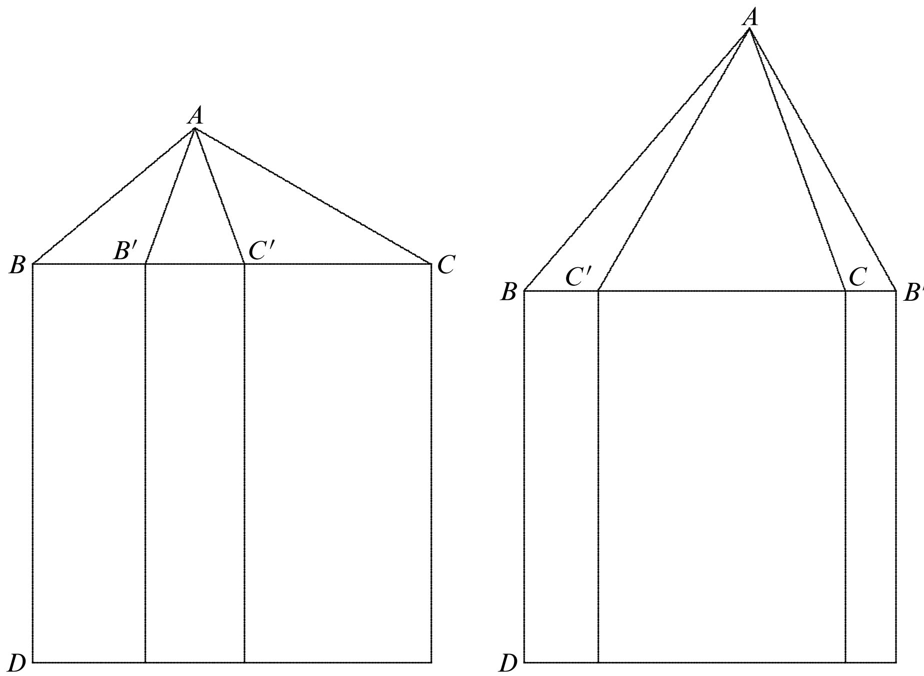
Answer. We shall consider only the case of a full rotation. The direct proportion of volumes and distances traveled by the centroid makes the extension to partial rotations an immediate extension. Consider then a right triangle of legs a and b rotated about an axis parallel to the leg of length b at a distance r from that leg. Assume first that the triangle lies on the side of the leg parallel to the axis opposite to the side containing the axis. This triangle generates the area outside a cylinder and inside a frustum of a cone. The cone has bottom radius $r+a$, top radius r and height b. The volume of the frustum is $\frac{\pi}{3}\bigl((a+r)^2(b+h)-r^2h\bigr) = \frac{\pi}{3}b(a^2+3ar+3r^2) = \frac{\pi}{3}a^2b + \pi abr + \pi br^2$. From this we subtract the volume of the cylinder inside, which is $\pi r^2 b$, getting a volume equal to $\frac{\pi}{3}ab(a+3r)$. Now the centroid M of this triangle is the intersection of its medians and lies at the vertex of a rectangle of sides $\frac{a}{3}$ and $\frac{b}{3}$ having one corner at the right angle of the triangle. In particular, its distance from the axis of rotation is $r+\frac{a}{3}$, and it therefore describes a circle of length $2\pi\bigl(r+\frac{a}{3}\bigr)$. Since the area is $\frac{1}{2}ab$, the product of these two is $\pi ab\bigl(r+\frac{1}{3}a\bigr)$, exactly the volume generated. (See the left-hand side of the accompanying figure.)

In the case when the triangle lies on the same side of the leg parallel to the axis as the axis itself, we must subtract a frustum having radii $r-a$ and r from a cylinder of radius r, and the centroid lies at distance $r-\frac{a}{3}$ from the axis. The argument is the same, however. (See the right-hand side of the accompanying figure.)

Now consider two areas A and B rotated about an axis in such a way that the centroid of area A is at distance r from the axis and the centroid of area B is at distance s, producing volumes V and W respectively. Then the centroid of the combined area is at distance

$$t = \frac{Ar}{A+B} + \frac{Bs}{A+B}$$

from the axis. The product of the area and the distance this centroid travels is $2\pi(A+B)t = 2\pi(Ar+Bs) = V+W$, which is exactly the volume produced. Now by induction on the number of regions, it follows that this theorem is true for any finite union of regions if it is true for each region individually. That includes all polygonal regions, since they can be decomposed into finite unions of triangles of the specified type.



$BD = BC$

Left: Thabit's theorem for an obtuse triangle. Right: Thabit's theorem when BC is not the longest side. In both cases $\overline{BD}(\overline{BB'} + \overline{CC'}) = \overline{AB}^2 + \overline{AC}^2$.

Finally, by the method of exhaustion, it follows that the theorem holds for all planar regions bounded by rectifiable curves.

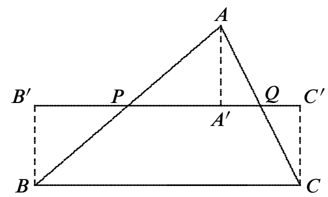
11.5. Explain how Thabit ibn-Qurra's generalization of the Pythagorean theorem reduces to that theorem when angle A is a right angle. What does the figure look like if angle A is obtuse? Is there an analogous theorem if BC is not the longest side of the triangle?

Answer. If angle A is a right angle, then $B' = C'$, the two rectangles built on side BC have a side in common, and their union is precisely the square on side BC. Thus the theorem becomes the Pythagorean theorem in that case.

If angle A is obtuse, there is a gap in the middle of side BC, but it remains true that the squares on AB and AC equal the corresponding rectangles built on parts of side BC. (See the left-hand side of the accompanying figure.)

To get a corresponding theorem when BC is not the longest side, you have to extend the side and let the angle constructed at A be larger than angle A. But a theorem can still be obtained. (See the right-hand side of the accompanying figure.)

11.6. One form of non-Euclidean geometry, known as doubly elliptic geometry, is formed by replacing the plane with a sphere and straight lines with great circles, that is, the intersections of the sphere with planes passing through its center. Let one "line" (great circle) be the equator of the sphere. Describe the equidistant curve generated by the endpoint of a "line segment" (arc of a great circle) of fixed length and perpendicular to the equator when the other endpoint moves along the equator. Why is this curve not a "line"?



Answer. The equidistant curve is what we call a parallel of latitude. It is a circle (actually a pair of circles, since the word "point" really should be interpreted as a pair of antipodal points). But it is not a great circle. Hence it is not a shortest path between any two of its points. Therefore it cannot be interpreted as a line.

11.7. Al-Haytham's attempted proof of the parallel postulate is fallacious because in non-Euclidean geometry two straight lines cannot be equidistant at all points. Thus in a non-Euclidean space the two rails of a railroad cannot both be straight lines. Assuming Newton's laws of motion (an object that does not move in a straight line must be subject to some force), show that in a non-Euclidean universe one of the wheels in a pair of opposite wheels on a train must be subject to some unbalanced force at all times. [Note: The spherical earth that we live on happens to be non-Euclidean. Therefore the pairs of opposite wheels on a train cannot both be moving in a great circle on the earth's surface.]

Answer. This problem is simply the remark that an object not moving along a geodesic must be subject to an unbalanced force. It shows the difficulty with the concept of a rigid body in non-Euclidean geometry. In special relativity there can be no rigid bodies. (If there were, it would be possible to transmit information instantaneously by moving a rigid bar in one place and having the other end of it write out information in another place.)

11.8. Prove that in any geometry, if a line passes through the midpoint of side AB of triangle ABC and is perpendicular to the perpendicular bisector of the side BC, then it also passes through the midpoint of AC. (This is easier than it looks: Consider the line that *does* pass through both midpoints, and show that it is perpendicular to the perpendicular bisector of BC; then argue that there is only one line passing through the midpoint of BC that is perpendicular to the perpendicular bisector of BC.)

Answer. If we draw the line through the midpoints P and Q of AB and AC respectively and drop perpendiculars AR, BS, CT to this line from A, B and C, we get two pairs of congruent triangles (by the angle-angle-side criterion), namely APR, BPS and AQR, CQT. It follows that AR, BS, and CT are all three equal. In particular, $BSTC$ is a Thabit quadrilateral. Since the perpendicular bisector of the base and summit are the same, it follows that the line $SPQT$ (which passes through the two midpoints P and Q) is perpendicular to the perpendicular bisector of BC. But there is only one perpendicular from P to this line, and hence the line through P perpendicular to the perpendicular bisector of BC must pass through Q, which was to be shown.

11.9. Use the previous result to prove, independently of the parallel postulate, that the line joining the midpoints of the lateral sides of a Thabit (Saccheri) bisects both diagonals.

Answer. The line joining the midpoints of the lateral sides of the Saccheri quadrilateral forms a second Saccheri quadrilateral on the same base. In particular, that line is perpendicular to the perpendicular bisector of the base. Therefore it passes through the midpoint of one side of the triangle formed by the base, one lateral side, and the diagonal. It must therefore pass through the midpoint of the diagonal.

CHAPTER 12

Modern Geometries

12.1. Judging from Descartes' remarks on mechanically drawn curves, should he have admitted the conchoid of Nicomedes among the legitimate curves of geometry?

Answer. Yes. Since a mechanical device exists for drawing it, it is sufficiently determinate to meet his standards.

12.2. Prove Menelaus' theorem and its converse. What happens if the points E and F are such that $AD : AE :: BD : BE$? (Euclid gave the answer to this question.)

Answer. The similarity of triangles EGF and CBF gives the following implication, where it is only necessary to add 1 to both sides of the equation.

$$\frac{FG}{BF} = \frac{FE}{CF} \Rightarrow \frac{BG}{BF} = \frac{CE}{CF}\,.$$

The similarity of triangles ADB and EDG similarly gives the following sequence of three implications.

$$\frac{BD}{DG} = \frac{AD}{DE} \Rightarrow \frac{BG}{DG} = \frac{AE}{DE} \Rightarrow \frac{DG}{BG} = \frac{DE}{AE} \Rightarrow \frac{BD}{BG} = \frac{AD}{AE}\,.$$

Here the first implication results from subtracting 1 from both sides, the second from taking reciprocals, and the third from adding 1 to both sides. Now we simply multiply the last two equalities to get

$$\frac{BD}{BF} = \frac{AD}{AE} \cdot \frac{CE}{CF}\,.$$

For the converse, we need to show that if the points A, B, D, E, and F are given and $AD : AE < BD : BF$, then the line through A and B meets the line through E and F on the side of the line through A and E on which B lies. But that is a simple consequence of the parallel postulate. If G is located on BD so that $AD : AE = BD : BG$, then it folows that $BF < BG$. Since EG is parallel to AB, the angles BAD and AEF total less than two right angles, and hence the lines AB and EF meet on that side of the transversal

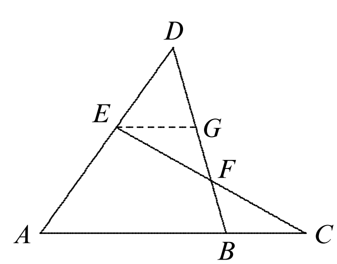


FIGURE 6. Menelaus' theorem for a plane triangle.

AD. The point of intersection C does satisfy the equality, as just proved, and since the ratio $CE : CF$ is strictly decreasing as C moves away from F, there can be only one point on the line EF having this property.

12.3. Use Menelaus' theorem to prove that two medians of a triangle intersect in a point that divides each in the ratio of 1:2.

Answer. This is the case where $AE = ED$ and $AB = BC$, taking the triangle to be ACD. Since $AD = 2AE$ in this case, we get

$$\frac{BD}{BF} = 2\frac{CE}{CF}.$$

By symmetry,

$$\frac{EC}{EF} = 2\frac{DB}{DF},$$

which we rewrite as

$$\frac{EF}{EC} = \frac{1}{2}\cdot\frac{DF}{DB}.$$

Subtracting both sides of this equality from 1 gives

$$\frac{CF}{CE} = \frac{1}{2} + \frac{1}{2}\cdot\frac{BF}{DB}.$$

Now multiplying this equality by the first equality, we get

$$\frac{1}{2}\cdot\frac{BD}{BF} + \frac{1}{2} = 2,$$

which can be rewritten as

$$\frac{BD}{BF} = 3.$$

12.4. Deduce Brianchon's theorem for a general conic from the special case of a circle. How do you interpret the case of a regular hexagon inscribed in a circle?

Answer. Consider a hexagon inscribed in a conic section, where the conic section is precisely that: the intersection of a plane and a cone. Each edge of the hexagon determines a unique plane through the vertex of the cone. Taking a fixed plane perpendicular to the axis of the cone, so that its section is a circle, we see that the planes determined by the sides of the hexagon intersect this plane in a hexagon inscribed in the circle. It is the projection of the hexagon inscribed in the original conic onto the plane of the circle. Conversely, the original hexagon is the projection of the other hexagon onto the plane of the original conic section. Since the opposite edges of the latter, when extended, meet in three collinear points and projections preserve collinearity, we see that the same must be true for the opposite edges of the original hexagon.

The three points of intersection of three pairs of parallel lines all lie on the same "line at infinity." If there is only one pair of parallel opposite edges, the line determined by the intersections of the other two pairs of opposite edges will be parallel to the pair of parallel edges. Hence the point at infinity where the other pair of edges meets lies on the line determined by the points where the other two pairs meet.

12.5. Fill in the details of Plücker's proof of Brianchon's theorem, as follows: Suppose that the equation of the conic is $q(x, y) = y^2 + r_1(x)y + r_2(x) = 0$, where $r_1(x)$ is a linear polynomial and $r_2(x)$ is quadratic. Choose coordinate axes not parallel to any of the sides of the inscribed hexagon and such that the x-coordinates of all of its vertices will be different, and also choose the seventh point to have x-coordinate different from those of the six vertices. Then suppose that the polynomial generated by the three lines is $s(x, y) = y^3 + t_1(x)y^2 + t_2(x)y + t_3(x) = 0$, where $t_j(x)$ is of degree j, $j = 1, 2, 3$. Then there are polynomials $u_j(x)$ of degree j, $j = 1, 2, 3$, such that[1]

$$s(x, y) = q(x, y)\big(y - u_1(x)\big) + (u_2(x)y + u_3(x)) .$$

We need to show that $u_2 \equiv 0$ and $u_3 \equiv 0$. At the seven points on the conic where both $q(x, y)$ and $s(x, y)$ vanish it must also be true that $u_2(x)y + u_3(x) = 0$. Rewrite the equation $q(x, y) = 0$ at these seven points as

$$(u_2y)^2 + r_1u_2(u_2y) + u_2^2r_2 = 0$$

observe that at these seven points $u_2y = -u_3$, so that the polynomial $u_3^2 - r_1u_2u_3 + u_2^2r_2$, which is of degree 6, has seven distinct zeros. It must therefore vanish identically, and that means that

$$(2u_3 - r_1u_2)^2 = u_2^2(r_1^2 - 4r_2) .$$

This means that either u_2 is identically zero, which implies that u_3 also vanishes identically, or else u_2 divides u_3. Prove that in the second case the conic must be a pair of lines, and give a separate argument in that case.

Answer. Most of the work is already done. In the case when $u_2(x)$ divides $u_3(x)$, say $u_3(x) = u_2(x)l(x)$, where $l(x)$ is a linear polynomial of the form $fx + g$, we can divide u_2^2 out of the equation, getting $l^2 - r_1l + r_2 = 0$, and then

$$q(x, y) = \big(y + l(x)\big)\big(y - l(x) + r_1(x)\big) .$$

Thus the conic is a pair of lines, either parallel or intersecting, and the inscribed hexagon has some parallel sides. It may even happen that two of the points of intersection coincide in this case. But in any case, this situation lies within the realm of the straight line, that is, the kind of geometry Euclid would have considered obvious.

12.6. Consider the two equations

$$\begin{aligned} xy &= 0, \\ x(y-1) &= 0. \end{aligned}$$

Show that these two equations are independent, yet have infinitely many common solutions. What kind of conic sections do these equations represent?

Answer. They are independent since the assumption that both are true leads to the conclusion $x = 0$. But the converse of that statement is also true, and hence the equations have infinitely many common solutions. Each represents a pair of intersecting lines, which is a degenerate hyperbola.

12.7. Consider the general cubic equation

$$Ax^3 + Bx^2y + Cxy^2 + Dy^3 + Ex^2 + Fxy + Gy^2 + Hx + Iy + J = 0,$$

which has 10 coefficients. Show that if this equation is to hold for the 10 points $(1, 0)$, $(2, 0)$, $(3, 0)$, $(4, 0)$, $(0, 1)$, $(0, 2)$, $(0, 3)$, $(1, 1)$, $(2, 2)$, $(1, -1)$, all 10 coefficients $A, \ldots, J$ must be zero. In general, then, it is not possible to pass a curve of degree 3 through any 10 points in the plane. Use linear algebra to show that it is always possible to pass a curve of degree 3 through any nine points, and that the curve is generally unique.

On the other hand, two *different* curves of degree 3 generally intersect in 9 points, a result known as Bézout's theorem after Étienne Bézout (1730–1783), who stated it around

[1] Note that $u_1(x)$ in the equation given in the text has been corrected here to $y - u_1(x)$.

1758, although Maclaurin had stated it earlier. How does it happen that while nine points generally determine a *unique* cubic curve, yet *two distinct* cubic curves generally intersect in nine points? [*Hint*: Suppose that a set of eight points $\{(x_j, y_j) : j = 1, \ldots, 8\}$ is given for which the system of equations for $A, \ldots, J$ has rank 8. Although the system of linear equations for the coefficients is generally of rank 9 if another point is adjoined to this set, there generally is a point (x_9, y_9), the ninth point of intersection of two cubic curves through the other eight points, for which the rank will remain at 8.]

Answer. The ten equations that result from putting in these ten values of (x, y) are

$$\begin{aligned} A + E + H + J &= 0 \\ 8A + 4E + 2H + J &= 0 \\ 27A + 9E + 3H + J &= 0 \\ 64A + 16E + 4H + J &= 0 \\ D + G + I + J &= 0 \\ 2D + 4G + 8I + J &= 0 \\ 3D + 9G + 27I + J &= 0 \\ A + B + D + D + E + F + G + H + I + J &= 0 \\ 8A + 8B + 8C + 8D + 4E + 4F + 4G + 2H + 2I + J &= 0 \\ A - B + C - D + E - F + G + H - I + J &= 0\,. \end{aligned}$$

The first four of these are a system whose determinant (the Vandermonde determinant) is not zero. Hence $A = E = H = J = 0$. Given that $J = 0$, the next three equations are likewise a system whose (Vandermonde) determinant is not zero, and so $D = G = I = 0$. Finally, the last three equations now say

$$\begin{aligned} B + C + F &= 0 \\ 8B + 8C + 4F &= 0 \\ -B + C - F &= 0\,, \end{aligned}$$

and it is easy to see that these equations imply $B = C = F = 0$.

On the other hand, for any nine points whatsoever, the corresponding set of homogeneous linear equations has more unknowns than equations and therefore has nonzero solutions.

12.8. Find the Gaussian curvature of the hyperbolic paraboloid $z = (x^2 - y^2)/a$ at each point using x and y as parameters.

Answer. Let us take advantage of what we know about vectors to explain again in brief terms how the curvature is computed. With the parameterization $(p, q) \mapsto \mathbf{r}(p, q)$, we compute a normal vector $\mathbf{n}(p, q) = \frac{\partial \mathbf{r}}{\partial p} \times \frac{\partial \mathbf{r}}{\partial q}$. The surface then has an "element of area" $dA = |\mathbf{n}|\, dp\, dq$. The spherical map $\rho(p, q) = \frac{1}{|\mathbf{n}(p,q)|}\mathbf{n}(p, q)$ is then a mapping into the unit sphere and has its own normal vector $\mathbf{N}(p, q) = c(p, q)\rho(p, q)$, where $c(p, q)$ is a scalar-valued function. The reason $\mathbf{N}$ has this form is that the normal line to a sphere at any point passes through the center of the sphere. Hence the element of area for the spherical map is $dS = |c(p, q)|\, dp\, dq$. The curvature is the ratio

$$\frac{c(p, q)}{|\mathbf{n}(p, q)|}\,.$$

In particular, its sign depends entirely on the sign of $c(p, q)$.

Since we are taking $x = p$, $y = q$, $z = (p^2 - q^2)/a$, we find that the element of the area of the surface is

$$dA = \frac{1}{a}\sqrt{a^2 + 4(p^2 + q^2)}\,dp\,dq = \frac{\Delta}{a}\,dp\,dq\,,$$

where $\Delta = \sqrt{a^2 + 4(p^2 + q^2)}$.

The normal line to this surface has direction numbers $(-2p/a, 2q/a, 1)$, so that the corresponding spherical map takes (p, q) to the point

$$\rho(p,q) = \left(\frac{-2p}{\sqrt{a^2 + 4(p^2 + q^2)}}, \frac{2q}{\sqrt{a^2 + 4(p^2 + q^2)}}, \frac{a}{\sqrt{a^2 + 4(p^2 + q^2)}}\right).$$

Although it is messy, one can compute that

$$\mathbf{N}(p,q) = -\frac{4a}{\Delta^3}\rho(p,q)\,.$$

Therefore the curvature is

$$-\frac{4a^2}{\Delta^4} = \frac{-4a^2}{(a^2 + 4(p^2 + q^2))^2}\,.$$

12.9. Find the Gaussian curvature of the pseudosphere obtained by revolving a tractrix about the x-axis. Its parameterization can be taken as[2]

$$\mathbf{r}(u,v) = \left(u - a\tanh\left(\frac{u}{a}\right), a\cos(v)\operatorname{sech}\left(\frac{u}{a}\right), a\sin(v)\operatorname{sech}\left(\frac{u}{a}\right)\right).$$

Observe that the elements of area on both the pseudosphere and its map to the sphere vanish when $u = 0$. (In terms of the first and second fundamental forms, $E = 0 = g$ when $u = 0$.) Hence curvature is undefined along the circle that is the image of that portion of the parameter space. Explain why the pseudosphere can be thought of as "a sphere of imaginary radius." Notice that it has a cusp along the circle in which it intersects the plane $x = 0$.

Answer. We shall confine ourselves to the range $u > 0$. The identites for trigonometric and hyperbolic functions yield a normal vector

$$\mathbf{n}(u,v) = -a\tanh\left(\frac{u}{a}\right)\operatorname{sech}\left(\frac{u}{a}\right)\left(\operatorname{sech}\left(\frac{u}{a}\right), \cos(v)\tanh\left(\frac{u}{a}\right), \sin(v)\tanh\left(\frac{u}{a}\right)\right).$$

The "vector" factor in this expression happens to *be* a unit vector. Hence the element of surface area on the pseudosphere is

$$a\tanh\left(\frac{u}{a}\right)\operatorname{sech}\left(\frac{u}{a}\right),$$

and the spherical mapping is

$$\rho(u,v) = -\left(\operatorname{sech}\left(\frac{u}{a}\right), \cos(v)\tanh\left(\frac{u}{a}\right), \sin(v)\tanh\left(\frac{u}{a}\right)\right).$$

We then find quite easily that

$$\mathbf{N}(u,v) = -\frac{1}{a}\tanh\left(\frac{u}{a}\right)\operatorname{sech}\left(\frac{u}{a}\right)\rho(u,v)\,.$$

Hence the curvature of the pseudosphere is $-\frac{1}{a^2}$. It represents a "sphere of radius $a\sqrt{-1}$."

[2] This formula is emended from the less general and more messy one in the text.

12.10. Prove that the Euler relation $V - E + F = 2$ for a closed polyhedron is equivalent to the statement that the sum of the angles at all the vertices is $(2V - 4)\pi$, where V is the number of vertices. [*Hint:* Assume that the polygon has F faces, and that the numbers of edges on the faces are $e_1, \ldots, e_F$. Then the number of edges in the polyhedron is $E = (e_1 + \cdots + e_F)/2$, since each edge belongs to two faces. Observe that a point traversing a polygon changes direction by an amount equal to the exterior angle at each vertex. Since the point returns to its starting point after making a complete circuit, the sum of the *exterior* angles of a polygon is 2π. Since the interior angles are the supplements of the exterior angles, we see that their sum is $e_i\pi - 2\pi = (e_i - 2)\pi$. The sum of all the interior angles of the polyhedron is therefore $(2E - 2F)\pi$.]

Answer. The hint essentially takes all the work out of this. The sum of all the angles at all vertices will be $(2E-2F)\pi$. That will be equal to $(2V-4)\pi$ if and only if $V-E+F=2$.

12.11. Give an informal proof of the Euler relation $V - E + F = 2$ for closed polyhedra, assuming that every vertex is joined by a sequence of edges to every other vertex. [*Hint:* Imagine the polyedron inflated to become a sphere. That stretching will not change V, E, or F. Start drawing the edges on a sphere with a single vertex, so that $V = 1 = F$ and $E = 0$. Show that adding a new vertex by distinguishing an interior point of an edge as a new vertex, or by distinguishing an interior point of a face as a new vertex and joining it to an existing vertex, increases both V and E by 1 and leaves F unchanged, while drawing a diagonal of a face increases E and F by 1 and leaves V unchanged. Show that the entire polyhedron can be constructed by a sequence of such operations.]

Answer. As long as every vertex in the polyhedron belongs to some edge and every edge is connected to some other edge, one can draw all the edges this way. The operations described leave no isolated vertices or edges, when there is more than one vertex, and of course such isolated vertices or edges have to be ruled out in order to get the formula.

CHAPTER 13

Problems Leading to Algebra

13.1. What do the two problems of recovering two numbers from their sum and product or from their difference and product have to do with quadratic equations as we understand them today? Can we conclude that the Mesopotamians "did algebra"?

Answer. We have emphasized several times that the problem of finding two numbers given their sum (or difference) and product is equivalent to the problem of solving a quadratic equation. This fact is taught to high-school students as a way of solving quadratic equations whose roots are integers (or rational numbers), the method students commonly refer to as FOIL. That is, given a quadratic equation px^2+qx+r with integer coefficients, it is known that a rational root $\frac{m}{n}$ must be such that m divides r and n divides p. In that way, the list of possible rational roots is reduced to a small set and one can then by brute force search determine if there are any rational roots. The same principle applies to equations of any degree, and what we know as Galois theory sets out from this point. That is, this algorithm determines all rational roots that exist, and so one can restrict attention to equations that have no rational roots.

The Mesopotamians knew the basic principle that we would phrase as the formula $\left(\frac{a+b}{2}\right)^2 = \left(\frac{a-b}{2}\right)^2 + ab$ and so were able to get $(a+b)/2$ knowing $(a-b)/2$ and ab. Likewise they could get $(a-b)/2$ knowing $(a+b)/2$ and ab. Having $(a+b)/2$ and $(a-b)/2$, they could then easily get both a and b. Although their method was not *stated* as a formal rule in any of the tablets, it was *used* so consistently that there is no doubt it was *learned* as a rule. On that basis, I have no hesitation in saying that they did algebra.

13.2. You can verify that the solution of the problem from tablet AO 8862 (15 and 12) given by the author is not the only possible one. The numbers 14 and 13 will also satisfy the conditions of the problem. Why didn't the author give this solution?

Answer. Most likely the author wasn't trying to find *all* solutions, so one was enough. Possibly the use of the words for *length* and *width* required length to be larger than width. In that case, the method used, which involved adding 2 to the width, would have made the new width larger than the length.

13.3. Of what practical value are the problems we have called "algebra"? Taking just the quadratic equation as an example, the data can be construed as the area and the semiperimeter of a rectangle and the solutions as the sides of the rectangle. What need, if any, could there be for solving such a problem? Where are you ever given the perimeter and area of a rectangle and asked to find its shape?

Answer. While I wouldn't be dogmatic about this, I think one can say that this problem practically never arises in everyday life. If there had been a practical application, one would expect to find it somewhere in the cuneiform tablets or the Chinese and Hindu classics. But we don't. Those documents are full of problems about ladders leaning against walls,

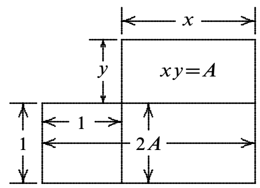


FIGURE 3. Another scenario to "fit" a text on cuneiform tablet AO 6670.

broken bamboo, and so on. Applications of algebra alone are not common, and those that do exist seem to have arisen after algebra was already in use. The algebra of economics and finance, input-output analysis, and so forth, are modern topics.

13.4. Figure 3 gives a scenario that can be fit to the data in AO 6670. Given a square 1 unit on a side, in the right angle opposite one of its corners construct a rectangle of prescribed area A that will be one-third of the completed gnomon. Explain how the figure fits the statement of the problem. (As in Section 2, this scenario is *not* being proposed as a serious explanation of the text.)

Answer. We are given a rectangle of unknown width $1 + x$ and height 1. We need to construct a second rectangle on top of it having width x one less than this width and area exactly half of it. That leads to the equations $xy = A$, $x+1 = 2A$, so that $y(2A-1) = A$, and so $y = \frac{A}{2A-1}$ and $x = \frac{1}{2A-1}$.

13.5. Given a cubic equation

$$ax^3 + bx^2 + cx = d\,,$$

where all coefficients are assumed positive, let $A = (2c^3)/(27d^2) + (bc)/(3d) + a$, $B = (c^2)/(3d) + b$, and $t = (3Adx)/(3Bd - cBx)$, that is, $x = (3Bdt)/(3Ad + cAt)$. Show that in terms of these new parameters, this equation is

$$t^3 + t^2 = \frac{A^2 d}{B^3}\,.$$

It could therefore be solved numerically by consulting a table of values of $t^3 + t^2$. [*Again a caution*: The fact that such a table exists and could be used this way does not imply that it *was* used this way, any more than the fact that a saucer can be used to hold paper clips implies that it was designed for that purpose.]

Answer. Fittingly, this problem is solved by tedious algebraic manipulation, unfortunately made more tedious than necessary—in fact, impossible—by misprints in the definition of A and the transformations between x and t (corrected here).

It is easier to take this one step at a time. Let us first set $x = 3du/(3d + cu)$, where u is a new variable, namely $u = 3dx/(3d-cx)$. Substituting this value of x in the equation, clearing the denominators, and canceling d gives

$$27ad^2u^3 + 9bdu^2(cu + 3d) + 3cu(cu + 3d)^2 = (cu + 3d)^3\,.$$

Gathering like terms here yields

$$Fu^3 + Gu^2 = H\,,$$

where $F = 27ad^2 + 9bcd + 2c^3$, $G = 27bd^2 + 9c^2d$, and $H = 27d^3$. The point is that the parameters in the change of variable were chosen so as to make the linear term in

u drop out. If we multiply this equation by F^2/G^3, we get

$$\frac{F^3}{G^3}u^3 + \frac{F^2}{G^2}u^2 = HF^2/G^3 \,.$$

Then, if we define a new variable t as $t = Fu/G$ we get the equation

$$t^3 + t^2 = K \,,$$

where $K = HF^2/G^3$. Observe that $F = 27d^2A$ and $G = 27d^2B$, so that $HF^2/G^3 = A^2d/B^3$, and $t = Au/B = 3Adx/(3Bd - cBx)$.

13.6. Considering the origin of algebra in the mathematical traditions we have studied, do you find a point in their development at which mathematics ceases to be a disjointed collection of techniques and becomes systematic? What criteria would you use for defining such a point, and where would you place it in the mathematics of Egypt, Mesopotamia, Greece, China, and India?

Answer. Like all periodizations in history, this one is an artificial boundary. Still, I have always been intrigued by Gillispie's notion of an "edge of objectivity" in the physical sciences, an edge not yet reached in the social sciences and maybe unreachable there. I would like to define such an edge in mathematics, corresponding to the progression from directly applicable mathematics to mathematics studied for its own sake. This point would have to be different in different areas of mathematics. Considering the thematic division of the present book, I would put two such divisions into the study of numbers and one into the study of space. These would correspond to the steps from counting to calculation and from calculation to number theory, and to the step from measurement to the abstract study of proportion in geometry. There are of course other significant quantitative steps, not all taken in serial order. All the vigorous areas of mathematics are continually getting infusions of new ideas from both outside and inside the subject itself.

CHAPTER 14

Equations and Algorithms

14.1. Problem 6 of Book 1 of the *Arithmetica* is *to separate a given number into two numbers such that a given fraction of the first exceeds a given fraction of the other by a given number.* In our terms this is a problem in two unknowns x and y, and there are four bits of data: the sum of the two numbers, which we denote by a, the two proper fractions r and s, and the amount b by which rx exceeds sy. Write down and solve the two equations that this problem involves. Under what conditions will the solutions be positive rational numbers (assuming that a, b, r, and s are positive rational numbers)? Compare your statement of this condition with Diophantus' condition, stated in very complicated language: *The last given number must be less than that which arises when that fraction of the first number is taken which exceeds the other fraction.*

Answer. The equations are

$$\begin{aligned} x + y &= a\,, \\ rx - sy &= b\,. \end{aligned}$$

The solution in our terms is

$$\begin{aligned} x &= \frac{sa + b}{r + s}\,, \\ y &= \frac{ra - b}{r + s}\,. \end{aligned}$$

These are certainly rational numbers, and x is positive. For $y > 0$ we need $ra > b$. Here a is the "first number," b is the "last given number," and the fraction is r.

14.2. Carry out the solution of the bundles of wheat problem from the *Jiu Zhang Suanshu*. Is it possible to solve this problem without the use of negative numbers?

Answer. We need to solve the equations

$$\begin{array}{ccccccc} 2x & + & y & & & = & 1 \\ & & 3y & + & z & = & 1 \\ x & & & + & 4z & = & 1\,. \end{array}$$

To avoid the use of negative numbers, one would have to use the slightly cumbersome "solve and substitute" approach: The first equation gives $y = 1 - 2x$. The second equation then gives $z = 1 - 3y = 6x - 2$. The third equation then gives $1 = x + 4z = 25x - 8$. Thus $x = \frac{9}{25}$, $y = \frac{7}{25}$, and $z = \frac{4}{25}$. Strictly speaking, none of these quantities is negative. However, that fact is known only in retrospect. During the solution process it was necessary to think of $2x$ being subtracted from 1. Since the *value* of x is not known at that point, it is necessary to store the 1 and the -2 in different "memory cells," even if one merely stores the 2 with a note that it is to be subtracted. Logically such an operation amounts to storing -2.

14.3. Solve the equation for the diameter of a town considered by Li Rui. [*Hint:* Since $x = -3$ is an obvious solution, this equation can actually be written as $x^3 + 3x^2 = 972$.]

Answer. Of course we could use the cubic formula to solve this problem, but it makes more sense to look for rational solutions first. We can guess a solution by writing $x^2(x + 3) = 972 = 81 \cdot 12 = 9^2(9 + 3)$, so that $x = 9$.

14.4. Solve the following legacy problem from al-Khwarizmi's *Algebra*: *A woman dies and leaves her daughter, her mother, and her husband, and bequeaths to some person as much as the share of her mother and to another as much as one-ninth of her entire capital. Find the share of each person.* It was understood from legal principles that the mother's share would be $\frac{2}{13}$ and the husband's $\frac{3}{13}$.

Answer. Let the daughter's share be x. According to the problem,

$$x + \frac{2}{13} + \frac{3}{13} + \frac{2}{13} + \frac{1}{9} = 1,$$

so that the daughter receives a share equal to $\frac{41}{117}$ of the estate.

14.5. Solve the problem of Abu Kamil in the text.

Answer. The conditions of the problem give

$$\frac{50}{x + 3} = \frac{50}{x} - 3\frac{3}{4}.$$

Clearing out the denominators results in

$$200x = 200(x + 3) - 15x(x + 3),$$

so that

$$x^2 + 3x - 40 = 0.$$

The only positive solution of this equation is $x = 5$.

14.6. If you know some modern algebra, explain, by filling in the details of the following argument, why it is not surprising that Omar Khayyam's geometric solution of the cubic cannot be turned into an algebraic procedure. Consider a cubic equation with rational coefficients but no rational roots,[1] such as $x^3 + x^2 + x = 2$. By Omar Khayyam's method, this equation is replaced with the system $y(z + 1) = 2$, $z^2 = (y + 1)(2 - y)$, one obvious solution of which is $y = 2$, $z = 0$. The desired value of x is the y-coordinate of the other solution. The procedure for eliminating one variable between the two quadratic equations representing the hyperbola and circle is a rational one, involving only multiplication and addition. Since the coefficients of the two equations are rational, the result of the elimination will be a polynomial equation with rational coefficients. If the root is irrational, that polynomial will be divisible by the minimal polynomial for the root over the rational numbers. However, a cubic polynomial with rational coefficients but no rational roots is itself the minimal polynomial for all of its roots. Hence the elimination will only return the original problem.

[1] If the coefficients are rational, their denominators can be cleared. Then all rational roots will be found among the finite set of fractions whose numerators divide the constant term and whose denominators divide the leading coefficient. There is an obvious algorithm for finding these roots.

Answer. If ξ is an irrational number satisfying an equation with rational coefficients, then there is a unique polynomial $p(x) = x^n + a_1 x^{n-1} + \cdots + a_{n-1}x + a_n$ with rational numbers $a_1, \ldots, a_n$ such that $p(\xi) = 0$. This polynomial is called the *minimal polynomial* of ξ over the rational numbers. The set of all polynomials $q(x)$ with rational coefficients such that $q(x) = 0$ can be shown (easily) to consist of precisely the polynomials of the form $p(x)r(x)$, *where* $r(x)$ *also has rational coefficients*. Since the original cubic polynomial that had this root is such a $q(x)$, it must be a multiple of $p(x)$. But since this cubic has no rational roots, neither $p(x)$ nor $r(x)$ can be a nonzero linear polynomial. Hence $p(x)$ must be cubic and $r(x)$ constant. That is, $p(x)$ is simply $q(x)$ divided by its leading coefficient. (Note that if $p(x)$ had a rational root w, we could divide it by $x - w$ and get a polynomial of smaller degree that vanishes at ξ, contrary to hypothesis.)

14.7. Why did al-Khwarizmi include a complete discussion of the solution of quadratic equations in his treatise when he had no applications for them at all?

Answer. Once again we ask why people do mathematics, and especially why they do algebra. Mr. Hamming's comments (see Chapter 2) begin to appear more and more pertinent. There is not, in my opinion, a very large gap between the scholastic philosphers discussing the hierarchy of angels and mathematicians a few centuries later classifying cubic equations. The usefulness of the latter did not appear until modern physics found a use for solving equations. (Finding the eigenvalues of matrices is an eminently useful thing to do, and that requires solving equations.)

14.8. Contrast the modern Western solution of the Islamic legacy problem discussed in the text with the solution of al-Khwarizmi. Is one solution "fairer" than the other? Can mathematics make any contribution to deciding what is fair?

Answer. This is a very subject. I give here a very personal statement on the meaning of moral propositions, with which the reader is (naturally) free to disagree.

On the one hand one finds dogmatic people who think the universe has revealed what is right and wrong. They usually go on to add that they know what the revelation was and that people who disagree with them are making an objective mistake. Among these people, surprisingly, one sometimes finds people who think that revelation amounts to complete relativism: that all ethical systems are equally good and no one should prefer one to another.

What is *objectively* clear when person A says that procedure X is fair and procedure Y is unfair is that person A wants to live in a world where X occurs and Y doesn't.[2] That (A's preference) is an objective fact. If B now says that X is unfair and Y is fair, we can make the same inferences about B. Then A and B are in *conflict* since they cannot both get what they want, but there is no *logical* or *factual* disagreement between them.

But one despairs of getting this understood. Inevitably the dogmatists will insist that a person who doesn't believe the universe endorses his moral system doesn't "really" have morals, that is, doesn't really have any preferences for the kind of society he wishes to live in and (they usually add) sees no difference between feeding the hungry and committing murder. In fact, everybody except a few sociopaths do have such preferences. In every society in history, people have been "socialized" to the extent that these preferences are

---

[2] In most cases more objective facts than that are available; A might, for instance, allege a direct divine revelation of his principles. We could then infer that A believes his principles have a source beyond his own preference. That is not the same as inferring that they really *do* have such a source, but A's belief in it is an objective fact nevertheless.

among the strongest motivations for action. Most people greatly prefer "death before dishonor." The points where A and B disagree about their preferences are the beginnings of politics.

As for the question posed here, I don't think mathematics has anything to contribute to the formation of moral principles. It may be useful in the practical implementation of a social or political program, of course. The specific application we are discussing seems to me one of the less crucial points of cultural difference, like the appropriate level of taxation that we discussed in connection with problems in the *Jiu Zhang Suan Shu*.

14.9. Consider the cubic equation of Sharaf al-Tusi's third type, which we write as $-x^3 - ax^2 + bx - c = 0$. Using Horner's method, as described in Section 2, show that if the first approximation is $x = m$, where m satisfies $3m^2 + 2am - b = 0$, then the equation to be satisfied at the second approximation is $y^2 - (3m + a)y^2 - (m^3 + am^2 - bm + c) = 0$. That is, carry out the algorithm for reduction and show that the process is

$$\begin{array}{c} -c \\ b \\ -a \\ -1 \end{array} \longrightarrow \begin{array}{c} -m^3 - am^2 + bm - c \\ -3m^2 - 2am + b\ (=0) \\ -3m - a \\ -1 \end{array}.$$

Answer. All we have to do is fill in the intermediate steps of the computation. They are as shown here:

$$\begin{array}{c} -c \\ b \\ -a \\ -1 \end{array} \longrightarrow \begin{array}{c} -c + bm - am^2 - m^3 \\ b - am - m^2 \\ -a - m \\ -1 \end{array} \longrightarrow \begin{array}{c} -c + bm - am^2 - m^3 \\ b - 2am - 3m^2 \\ -a - 2m \\ -1 \end{array} \longrightarrow \begin{array}{c} -c + bm - am^2 - m^3 \\ b - 2am - 3m^2 \\ -a - 3m \\ -1 \end{array}.$$

14.10. Consider Problem 27 of Book 1 of *De numeris datis*: *Two numbers are given whose sum is 10. If one is divided by 4 and the other by 2, the product of the quotients is 2. What are the two numbers?* Solve this problem in your own way, then solve it following Jordanus' recipe, which we paraphrase as follows. Let the two numbers be x and y, and let the quotients be e and f when x and y are divided by c and d respectively; let the product of the quotients be $ef = b$. Let $bc = h$, which is the same as fce or fx. Then multiply d by h to produce j, which is the same as xdf or xy. Since we now know both $x + y$ and xy, we can find x and y.

Answer. The statement about the quotients is a roundabout way of saying that the product of the two numbers is 16. We are thus looking for two numbers whose sum is 10 and whose product is 16. The numbers are therefore 2 and 8.

Jordanus' solution gives a general rule for solving problems of this sort, analogous to a formula in which numerical values can be substituted. He tends to invent a new letter to mark each stage of the solution process, carefully stating how it is related to all the letters encountered at previous stages, resulting in a total of 9 letters to describe a process with only two unknowns and four bits of data.

14.11. Solve the equation $x^3 + 60x = 992$ using the recipe given by Tartaglia.

Answer. We want two numbers whose difference is 992 and whose product is $\left(\frac{60}{3}\right)^3 = 8000$. Following the ancient method from Mesopotamia, we know that the average of the two numbers is the square root of the sum of their product and the square of half their difference, that is, the sum is $\sqrt{8000 + (496)^2} = 504$. the two numbers are therefore

$504+496 = 1000$ and $504-496 = 8$. Hence the solution is $\sqrt[3]{1000}-\sqrt[3]{8} = 10-2 = 8$. It is easily verified that this solution is correct.

14.12. How can you *prove* that $\sqrt[3]{\sqrt{108}+10}-\sqrt[3]{\sqrt{108}-10} = 2$?

Answer. A rigorous proof requires showing that the strictly increasing function $f(x) = x^3+6x$ takes the same value at both numbers. In other words, one can only get a useful form (namely 2) for the solution of the equation by proving that useful form satisfies the equation. (If you ask a hand calculator to do these computations, chances are it will tell you that the difference of the two cube roots has the value 2.000000001.)

14.13. If you know the polar form of complex numbers $z = r\cos\theta + i\,r\sin\theta$, show that the problem of taking the cube root of a complex number is equivalent to solving two of the classical problems of antiquity simultaneously, just as Viète claimed: the problem of two mean proportionals and the problem of trisecting the angle.

Answer. The complex number would be given by specifying the nonnegative number r (its absolute value) and θ (its argument). It can then be verified from the formulas $\cos 3\alpha = 4\cos^3\alpha - 3\cos\alpha$ and $\sin 3\alpha = \sin\alpha(4\cos^2\alpha - 1)$ (which are immediate consequences of the addition formulas for sine and cosine) that $\sqrt[3]{r}\Big[\Big(\cos\frac{\theta}{3} + i\sin\frac{\theta}{3}\Big)^3\Big] = z$

14.14. Consider Viète's problem of finding three numbers in direct proportion given the middle number and the difference between the largest and smallest. Show that this problem amounts to finding x and y given $\sqrt{xy}$ and $y-x$. How do you solve such a problem?

Answer. You are given that $y-x = d$ and that $x : a = a : y$, where you know d and $a = \sqrt{xy}$. If you square a, you get a known number, and now you know the product and difference of y and x. It has been known for thousands of years that finding x and y from this information is a problem having unique positive solutions.

14.15. Show that the equation $x^3 = px+q$, where $p > 0$ and $q > 0$, has the solution $x = \sqrt{4p/3}\cos\theta$, where $\theta = \frac{1}{3}\arccos\big((q\sqrt{27})/(2\sqrt{p^3})\big)$. In order for this inverse cosine to exist it is necessary and sufficient that $q^2/4 - p^3/27 \le 0$, which is precisely the condition under which the Cardano formula requires the cube root of a complex number. [*Hint*: Use the formula $4\cos^3\theta - 3\cos\theta = \cos(3\theta)$.]

Observe that

$$\theta = \frac{1}{3}\int_a^1 \frac{1}{\sqrt{1-t^2}}\,dt\,,$$

where $a = (q\sqrt{27})/(2\sqrt{p^3})$. Thus, the solution of the cubic equation has a connection with the integral of an algebraic function $1/y$, where y satisfies the quadratic equation $y^2 = 1-x^2$. This kind of connection turned out to be the key to the solution of higher-degree algebraic equations. As remarked in the text, Viète's solution of the cubic uses a transcendental method, even though an algebraic method exists.

Answer. Following the hint, all we have to do is substitute:

$$x^3 - px = \frac{p}{3}\sqrt{\frac{4p}{3}}\big(4\cos^3\theta - 3\cos\theta\big) = \frac{2\sqrt{p^3}}{\sqrt{27}}\cos(3\theta)\,.$$

But

$$\cos(3\theta) = \frac{q\sqrt{27}}{2\sqrt{p^3}}\,,$$

and so, $x^3 - px = q$.

CHAPTER 15

Modern Algebra

15.1. Prove that if every polynomial with real coefficients has a zero in the complex numbers, then the same is true of every polynomial with complex coefficients. To get started, let $p(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1}z + a_n$ be a polynomial with complex coefficients $a_1, \ldots, a_n$. Consider the polynomial $q(z)$ of degree $2n$ given by $q(z) = p(z)\overline{p(\bar{z})}$, where the overline indicates complex conjugation. This polynomial has real coefficients, and so by hypothesis has a complex zero z_0.

Answer. Since $q(z)$ has a zero z_0, it follows that $p(z_0)\overline{p(\bar{z}_0)} = 0$. Hence either $p(z_0) = 0$ or $p(\bar{z}_0) = 0$. In either case, $p(z)$ has a zero.

15.2. Formulate Cauchy's 1812 result as the following theorem and prove it: *Let p be a prime number, $3 \le p \le n$. If a subgroup of the symmetric group on n letters contains all permutations of order p, it is either the entire symmetric group or the alternating group.*

Answer. Before doing this, I would like to engage in a rant against the algebraists who write about permutations. It is not—to use a slightly outmoded colloquialism—"rocket science" to distinguish between a location and the occupant of that location. Therefore one has to wonder why algebraists have spent so many centuries confusing the two (and their readers) with their notation. To take just one example, from the classic textbook of Birkhoff and Mac Lane:[1] "The cyclic permutation $\gamma = (a_1, a_2, \ldots, a_n)$ carries a_i into a_{i+1}."

Indeed. What, pray tell, does "carried into" mean? Does it mean that the symbol a_i is *to occupy the position* formerly occupied by the symbol a_{i+1}? Or does it mean that the symbol a_i is *to be replaced* by the symbol a_{i+1}? Both of these assumptions are reasonable interpretations, but the two interpretations are exactly inverse to each other. After studying the examples of those writers thoughtful enough to provide any (that is by no means all of them) I usually conclude that they mean the former, but I wish they would say so more clearly.

It gets worse. If the notation (abc) is slightly ambiguous, things become doubly ambiguous when numbers are used instead of letters. Now one has to worry not only about whether symbols are replacing symbols or locations locations, but also whether the numbers refer to symbols or locations. Consider, for example, an arrangement of the numbers 1 through 5 such as $(2, 3, 5, 1, 4)$. Now ask yourself how the symbol (123) means to rearrange that.

(a) Is the symbol 1 to move to the position now occupied by 2, the symbol 2 to the position now occupied by 3, and the symbol 3 to the position now occupied by 1? If so, the result is $(1, 2, 5, 3, 4)$.

(b) Is the symbol 1 to be erased and replaced with 2, the symbol 2 erased and replaced with 3, and the symbol 3 erased and replaced by 1? If so, the result is $(3, 1, 5, 2, 4)$.

[1] *A Survey of Modern Algebra*, revised edition, Macmillan, New York, 1953, p. 136.

(c) Is the symbol in position 1 to be moved to position 2, the symbol in position 2 to position 3, and the symbol in position 3 to position 1? If so, the result is $(5, 1, 2, 3, 4)$.

(d) Is the symbol in position 1 to be replaced by the symbol in position 2, the symbol in position 2 by the symbol in position 3, and the symbol in position 3 by the symbol in position 1? If so, the result is $(2, 5, 1, 3, 4)$.

It gets yet worse. Algebraists tend to hate parentheses. That wouldn't be so bad if, like analysts, they wrote functional symbols to the left of their arguments; but many of them seem perversely to prefer the opposite. Where the analyst writes $\varphi(x)$, the algebraist prefers $x\varphi$. Very well, what does $(ab)(ac)$ mean? In which order do we perform these transpositions? In the textbooks I have read this notation means to start with $(a, b, c, \dots)$ and do first (ab), getting $(b, a, c, \dots)$, then (ac), getting $(b, c, a, \dots)$. Notice that if these transpositions were performed in the opposite order, the result would be the arrangement $(c, a, b, \dots)$.

End of rant. Let us return to the problem at hand, showing that a group of permutations that contains all permutations of order p, where p is an odd prime, is either the full symmetric group or the alternating group of even permutations. Just to be very clear: *Wherever the letters happen to be in an arrangment*, when the cycle (abc) is applied the result is that a goes to the position where b is, b goes to the position where c is, and c goes to the position where a is. Thus if we have, say (c, a, e, f, b), then the result of applying (abc) is the arrangement (b, c, e, f, a).

The trick is to use Cauchy's result that such a subgroup contains all 3-cycles. The proof of that fact for the case of $p = 5$ is quite simple: Just note that $(abc) = (adcbe)(aecdb)$. If there are only three letters to permute then the soubgroup certainly contains the alternating group, which consists precisely of the 3-cycles and the identity. Let us now assume that we are permuting more than 3 letters, so that the set of elements being permuted is $a, b, c, d, \dots$. Our subgroup contains all products of two interlocking transpositions, such as $(ab)(ac) = (acb)$, since these are 3-cycles. And since noninterlocking transpositions can also be written in this way, for example, $(ab)(cd) = (cad)(abc)$, it follows that every product of two transpositions is in the subgroup. Since every permutation is a product of transpositions, every even permutation is in the subgroup

Now assume the subgroup also contains one transposition, say (ab). Then it also contains $(ac) = (ab)(abc)$, $(ad) = (ac)(acd)$, and $(cd) = (ca)(cad)$. It thus contains all transpositions. Since every permutation is a product of transpositions, the subgroup contains all permutations whatsoever.

15.3. Cauchy's theorem that every cycle of order 3 can be written as the composition of two cycles of order m if $m > 3$ looks as if it ought to apply to cycles of order 2 also. What goes wrong when you try to prove this "theorem"?

Answer. If you look at the case of permutations of three objects, you can see that the cycles of order 3 cannot produce any cycle of order 2. The composition of either permutation of order 3 with itself is the other permutation of order 3. Their composition with each other is the identity permutation.

15.4. Let $S_j(a, b, c, d)$ be the jth elementary symmetric polynomial, that is, the sum of all products of j distinct factors chosen from $\{a, b, c, d\}$. Prove that $S_j(a, b, c, d) = S_j(b, c, d) + aS_{j-1}(b, c, d)$. Derive as a corollary that given a polynomial equation $x^4 - S_1(a, b, c, d)x^3 + S_2(a, b, c, d)x^2 + S_3(a, b, c, d)x + S_r(a, b, c, d) = 0 = x^4 - p_1x^3 + p_2x^2 - p_3x + p_4$ having a, b, c, d as roots, each elementary symmetric function in b, c, d

can be expressed in terms of a and the coefficients p_j: $S_1(b,c,d) = p_1 - a$, $S_2(b,c,d) = p_2 - aS_1(b,c,d) = p_2 - ap_1 + a^2$, $S_3(b,c,d) = p_3 - ap_2 + a^2 p_1 - a^3$.

Answer. Simply break the sum that is $S_j(a,b,c,d)$ into two smaller sums, one consisting of the terms that do not contain a as a factor (which must be $S_j(b,c,d)$—otherwise, the sum $S_j(a,b,c,d)$ wouldn't contain all possible terms) and those that do contain a as a factor (which must be $aS_{j-1}(b,c,d)$ for the same reason). If we define $S_0(b,c,d) = 1$, this formula holds with $j = 1$, and if we define $S_0(a,b,c,d) = 1$, $S_{-1}(b,c,d) = 0$, it also holds for $j = 0$. The equations given for expressing $S_j(b,c,d)$ as a polynomial in a with coefficients from among $p_1, \ldots, p_4$ are immediate, trivial computations. For example, $p_2 = S_2(a,b,c,d) = S_2(b,c,d) + aS_1(b,c,d) = S_2(b,c,d) + ap_1 - a^2$, and therefore $S_2(b,c,d) = p_2 - ap_1 + a^2$, and so on.

15.5. Prove that if z is a prime in the ring obtained by adjoining the pth roots of unity to the integers (where p is a prime), the equation

$$z^p = x^p + y^p$$

can hold only if $x = 0$ or $y = 0$.

Answer. This theorem fails for $p = 2$, since $5^2 = 3^2 + 4^2$. However, it can be restored by adjoining the fourth roots of unity, and in fact becomes trivial in that case. I shall discuss the "sufficiently typical" case of $p = 5$, leaving the easier case $p = 3$ for the reader to practice on. The general case is a little messier than I care to undertake. I am of course assuming that x, y, and z are integers.

Let us begin with some simple algebra. We note that the equation $z^5 - 1 = 0$ factors as $(z-1)(z^4 + z^3 + z^2 + z + 1) = 0$. Any root ω satisfying $\omega^4 + \omega^3 + \omega^2 + \omega + 1 = 0$, is a primitive 5th root of unity. You can easily verify that the other roots of $x^4 + x^3 + x^2 + x + 1 = 0$ are then ω^2, ω^3, and ω^4. (Since $\omega^5 = 1$, you have, for example, $(\omega^3)^4 + (\omega^3)^3 + (\omega^3)^2 + \omega^3 + 1 = \omega^2 + \omega^4 + \omega + \omega^3 + 1 = 0$. Hence ω^j is a primitive 5th root of unity for $j = 1, 2, 3, 4$.) Since the polynomial $f(x) = x^4 + x^3 + x^2 + x + 1$ is irreducible over the rational numbers, it is the minimal polynomial of ω. That is, if $g(x)$ is a polynomial with rational coefficients and $q(\omega) = 0$, then $g(x) = f(x)r(x)$ for some polynomial $r(x)$. In particular, if $g(x)$ has degree less than 5, then in fact $g(x) = rf(x)$ for some constant r, and all the coefficients of $g(x)$ are equal. Thus we have the proposition that $a\omega^4 + b\omega^3 + c\omega^2 + d\omega + e = 0$ with rational numbers a, b, c, d, e if and only if the four equations $a = b = c = d = e$ hold.

That being established, we note that since $p(x)$ is a monic polynomial (has leading coefficient 1) and $p(\omega^j) = 0$ for $j = 1, 2, 3, 4$, it follows that $x^4 + x^3 + x^2 + x + 1 = p(x) = (x-\omega)(x-\omega^2)(x-\omega^3)(x-\omega^4)$. Replacing x by $-x$ and multiplying the right-hand side by $1 = (-1)^4$, we get $x^4 - x^3 + x^2 - x + 1 = (x+\omega)(x+\omega^2)(x+\omega^3)(x+\omega^4)$. Multiplying both sides of this equation by $x + 1$, we get finally

$$x^5 + 1 = (x+1)(x+\omega)(x+\omega^2)(x+\omega^3)(x+\omega^4)\,.$$

As a corollary, taking $x = \frac{m}{n}$ and multiplying by n^5, we find that

$$m^5 + n^5 = (m+n)(m+\omega n)(m+\omega^2 n)(m+\omega^3 n)(m+\omega^4 n)\,.$$

What we have done so far extends easily to any odd prime number p. Now we use the special case $p = 5$ to save some messiness.

We wish to find all units in the ring $\mathbb{Z}[\omega]$ of the form $a + b\omega$, where a and b are integers. Obviously ± 1 and $\pm\omega$ are of this form. Since we have not said *which* of the five roots we have in mind, we have now identified ten units in the ring. We shall prove that there are precisely two others of the form $a + b\omega$, namely $\pm(1 + \omega)$. Again, since we have not specified ω, it follows that there are precisely two others of the form $a + b\omega^j$, $j =, 2, 3, 4$, namely $\pm(1+\omega^j)$. Suppose then that $a+b\omega$ is a unit. Let $w = p+q\omega+r\omega^2+s\omega^3+t\omega^4$ be such that $w(a + b\omega) - 1 = 0$. According to our criterion for this equality, we get the following four equations for p, q, r, s, and t.

$$\begin{array}{ccccccccccc} & & & & & & a(p-t) & + & b(t-s) & = & 1\,, \\ bp & + & (a-b)q & - & ar & & & & & = & 0\,, \\ & & bq & + & (a-b)r & - & as & & & = & 0\,, \\ & & & & br & + & (a-b)s & - & at & = & 0\,. \end{array}$$

The first of these equations tells us that a and b must be relatively prime and cannot both be zero. If one of them is zero, the other must be ± 1, and we have already considered this case. Therefore we assume that neither a nor b is zero. If we solve the remaining three equations to get p in terms of a, b, s, and t and substitute this value into the first equation, the result is

$$b^3 = (t - s)(a^4 - a^3 b + a^2 b^2 - ab^3 + b^4)\,.$$

Now let m be any prime divisor of b. Then, since a and b are relatively prime, m does not divide the second factor on the right, and hence m^3 must divide $t - s$. Since this is true for all prime divisors of b, we have $t - s = b^3 w$, and then

$$1 = w(a^4 - a^3 b + a^2 b^2 - ab^3 + b^4) = \frac{w(a^5 + b^5)}{a + b}\,.$$

The first equation here shows that $w = \pm 1$ and the second shows that $w > 0$, since $a^5 + b^5$ and $a + b$ obviously are of the same sign. We thus have

$$1 - a^2 b^2 = a^4 - a^3 b - ab^3 + b^4 = (a^3 - b^3)(a - b) = (a - b)^2(a^2 + ab + b^2) \ge 0\,.$$

It follows that $a^2 b^2 \le 1$ and so the only possible units that we have not considered have $a = \pm 1$, $b = \pm 1$. But this last equation can be satisfied in this case only if $a = b$, and so we have really only to consider the case $a = b = 1$. Our original equations for this case imply $p = r = t$ and $q = s = t-1$. That is, $(1+\omega)(t+(t-1)\omega+t\omega^2+(t-1)\omega^3+t\omega^4) = (1 + \omega)(-\omega - \omega^3)$, since $t(1 + \omega + \omega^2 + \omega^3 + \omega^4) = 0$. Multiplying this out now yields $-\omega - \omega^2 - \omega^3 - \omega^4 = 1$. We have now found the last two units of the form $a + b\omega$.

With these preliminaries out of the way, the proof can now proceed very efficiently. Assume that x, y, and z are positive integers such that z is prime in this ring and

$$z^5 = x^5 + y^5 = (x + y)(x + \omega y)(x + \omega^2 y)(x + \omega^3 y)(x + \omega^4 y)\,.$$

Now z divides z^5, and so it divides the product on the right. Being prime (as opposed to being merely irreducible), z must divide one of the factors on the right, say $x + \omega^k y = uz$. We can then cancel one factor of z from both sides and start over with z^4 on the left-hand side. Eventually we will arrive at an equation

$$z = (x + \omega^r y)w\,.$$

Since z is irreducible, either $x + \omega^r y$ is a unit in this ring or w is a unit. The first assumption, as we saw, implies that either $x = 0$ or $y = 0$ or $x = y$. Since there are no nontrivial solutions when $x = y$, we need only show that the second assumption also leads to the conclusion that $x = 0$ or $y = 0$. This time we add the assumption that $x \ne y$, so that $x + \omega^j y$ is not a unit unless $x = 0$ or $y = 0$.

If w is a unit, then we have $x + \omega^r y = vz$, where v is a unit. We can now go back to the original equation and cancel one factor of z, getting the equation

$$z^4 = v \prod_{j \neq r} (x + \omega^j y)\,.$$

We now continue the reduction once again, getting eventually $z = u(x + \omega^s y)$, where $s \neq r$. The same argument now implies that u is a unit. We can now start over once again with an equation

$$z^3 = vu \prod_{j \neq r,s} (x + \omega^j y)\,.$$

Continuing as long as necessary, we eventually find that $tz = x + y$, where t is a unit in the ring. If $t = t_0 + t_1\omega + t_2\omega^2 + t_3\omega^3 + t_4\omega^4$, this equation says that

$$0 = zt_0 - (x + y) + z(t_1\omega + t_2\omega^2 + t_3\omega^3 + t_4\omega^4)\,.$$

Our criterion for this equation to hold then implies that $t_1 = t_2 = t_3 = t_4$, and therefore

$$0 = z(t_0 - t_1) - (x + y)\,.$$

But since $t_0 - t_1$ must be a unit and x, y, and z are all positive, it follows that $z = x + y$. We then have the equation

$$(x + y)^5 = x^5 + y^5\,,$$

which obviously cannot hold unless either $x = 0$ or $y = 0$. This same proof works for any prime p, provided one can establish what the units of the form $a + b\omega$ are.

15.6. Consider the complex numbers of the form $z = m + n\omega$, where $\omega = -1/2 + \sqrt{-3}/2$ is a cube root of unity. Show that $N(z) = m^2 - mn + n^2$ has the property $N(zw) = N(z)N(w)$ and that $N(z+w) \le 2\big(N(z) + N(w)\big)$. Then show that a Euclidean algorithm exists for such complex numbers: Given z and $w \neq 0$, there exist q and r. Such that $z = qw + r$ where $N(r) < N(w)$. Thus, a Euclidean algorithm exists for these numbers, and so they must exhibit unique factorization. [*Hint*: $N(z) = |z|^2$. Show that for every complex number u there exists a number q of this form such that $|q - u| < 1$. Apply this fact with $u = z/w$ and *define* r to be $z - qw$.]

Answer. The identity $N(zw) = N(z)N(w)$ is trivial algebra, since it amounts to the well-known fact $|zw| = |z|\,|w|$ for complex numbers. As for $N(z + w) \le 2\big(N(z) + N(w)\big)$, we observe that $\sqrt{N(z + w)} \le \sqrt{N(z)} + \sqrt{N(w)}$, and squaring then gives the required result, since $2\sqrt{ab} \le a + b$.

Now let $u = x + iy$. Let n be the nearest integer to $2y/\sqrt{3}$, so that $\big|n - \frac{2y}{\sqrt{3}}\big| \le \frac{1}{2}$, and therefore $\big|\frac{n\sqrt{3}}{2} - y\big| < \frac{\sqrt{3}}{4}$. Then let m be the nearest integer to $\frac{n}{2} + x$, so that $\big|m - \frac{n}{2} - x\big| \le \frac{1}{2}$. We then have

$$|m + n\omega - u|^2 = \Big(m - \frac{n}{2} - x\Big)^2 + \Big(\frac{n\sqrt{3}}{2} - y\Big)^2 \le \frac{1}{4} + \frac{3}{16} < 1\,.$$

Now suppose u and w are any two numbers of this form with $w \neq 0$. Taking $u = z/w$ and defining $r = z - qw$, we get $|r|^2 = |w|^2|u - q|^2 < |w|^2$, that is, $N(r) < N(w)$. That this property implies a Euclidean algorithm and unique factorization was proved in Problem 8.10.

15.7. Show that in quaternions the equation $X^2+r^2 = 0$, where r is a positive real number (scalar), is satisfied precisely by the quaternions $X = x + \xi$ such that $x = 0$, $|\xi| = r$, that is, by all the points on the sphere of radius r. In other words, in quaternions the square roots of negative numbers are simply the nonzero vectors in three-dimensional space. Thus, even though quaternions act "almost" like the complex numbers, the absence of a commutative law makes a great difference when polynomial algebra is considered. A linear equation can have only one solution, but a quadratic equation can have an uncountable infinity of solutions.

Answer. This is a routine computation:

$$X^2 + r^2 = (x^2 - \xi \cdot \xi + r^2) + (2x\xi)\,.$$

If the vector part of this quaternion is to be zero, either $x = 0$ or $\xi = 0$. But $\xi = 0$ implies $x^2 + r^2 = 0$, which is impossible if $r \neq 0$. Hence we must have $x = 0$ and $|\xi| = r$.

CHAPTER 16

The Calculus

16.1. Show that the Madhava–Jyeshtadeva formula given at the beginning of the chapter is equivalent to

$$\theta = \sum_{k=0}^{\infty} (-1)^k \frac{\tan^{2k+1}\theta}{2k+1},$$

or, letting $x = \tan\theta$,

$$\arctan x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1}.$$

Answer. This is routine, given our notation. From the formula given in the textbook, simply divide by r, then use the fact that $\tan\theta = \frac{\sin\theta}{\cos\theta}$.

16.2. Consider an ellipse with semiaxes a and b and a circle of radius b, both circle and ellipse lying between a pair of parallel lines a distance $2b$ apart. For every line between the two lines and parallel to them, show that the portion inside the ellipse will be a/b times the portion inside the circle. Use this fact and Cavalieri's principle to compute the area of the ellipse. This result was given by Kepler.

Answer. Without any loss of generality, we can take the equation of the circle to be $x^2 + y^2 = b^2$ and the equation of the ellipse to be $b^2x^2 + a^2y^2 = a^2b^2$. Now consider any vertical line $y = t$, where $-b \le t \le b$. This line intersects the circle in the two points $x = \pm\sqrt{b^2 - t^2}$ and the ellipse in the two points $x = \pm(a/b)\sqrt{b^2 - t^2}$. Hence the ratio of the two lines is a/b, as asserted. It follows that the area of the ellipse is a/b times the area of the circle, that is $\pi b^2(a/b) = \pi ab$.

16.3. Show that the point at which the tangent to the curve $y = f(x)$ intersects the y axis is $y = f(x) - xf'(x)$, and verify that the area under this curve—more precisely, the integral of $f(x) - xf'(x)$ from $x = 0$ to $x = a$—is twice the area between the curve $y = f(x)$ and the line $ay = f(a)x$ between the points $(0, 0)$ and $(a, f(a))$. This result was used by Leibniz to illustrate the power of his infinitesimal methods.

Answer. The equation of the tangent in (s, t) coordinates is $t - f(x) = f'(x)(s - x)$. When we set $s = 0$, we find that $t = f(x) - xf'(x)$. Then

$$\int_0^a f(x) - xf'(x)\,dx = \int_0^a f(x)\,dx - \int_0^a xf'(x)\,dx.$$

Integrating by parts in this last expression we find

$$\int_0^a xf'(x)\,dx = xf(x)\Big|_0^a - \int_0^a f(x)\,dx = af(a) - \int_0^a f(x)\,dx.$$

Putting these results together, we find

$$\int_0^a f(x) - xf'(x)\,dx = 2\left(\int_0^a f(x)\,dx - \frac{1}{2}af(a)\right).$$

Since $\frac{1}{2}af(a)$ is the area under the line $y = f(a)x/a$ from $x = 0$ to $x = a$ (that is, it is the area of a triangle of base a and height $f(a)$), we are done.

16.4. Recall that Eudoxus solved the problem of incommensurables by changing the definition of proportion, or rather, *making* a definition to cover cases where no definition existed before. Newton's "theorem" asserting that quantities that approach each other continually (we would say monotonically) and become arbitrarily close to each other in a finite time must become equal in an infinite time assumes that one has a definition of equality at infinity. What is the definition of equality at infinity? Since we cannot *actually* reach infinity, the definition will have to be stated as a potential infinity, that is, a statement about all possible finite times. Formulate the definition and compare Newton's solution of this difficulty with Eudoxus' solution of the problem of incommensurables.

Answer. If we formulate "equality after an infinite time" so as to avoid the actually infinite, we find ourselves saying that for any prescribed difference there is a finite time after which the quantities will differ by less than that amount. In that respect, Newton's "proposition" becomes a mere tautology. It says that quantities that become arbitrarily close to each other in a finite time must come closer than any prescribed difference in some finite time. Thus Newton's solution of the difficulty of "indivisibles" is, like Eudoxus' solution of the difficulty of incommensurables, an attempt to make a definition that fits intuition. Newton's attempt to turn his definition into a theorem resembles in many respects Euclid's attempt to define the term *point*. Hidden in his use of the phrase "after an infinite time" was the assumption that everyone had this intuitive notion and that one could intuit what happens at infinity from what happens at all finite times.

As Zeno showed, such is not the case if the infinity is "too big," as in the case of a line made up of points. You can't get anywhere near the properties of a line by considering large numbers of points. Eudoxus similarly gave an intuitive formulation of what it meant to compare two incommensurable magnitudes, essentially defining the difficulty away.

16.5. Draw a square and one of its diagonals. Then draw a very fine "staircase" by connecting short horizontal and vertical line segments in alternation, each segment crossing the diagonal. The total length of the horizontal segments is the same as the side of the square, and the same is true of the vertical segments. Now in a certain intuitive sense these segments approximate the diagonal of the square, seeming to imply that the diagonal of a square equals twice its side, which is absurd. Does this argument show that the method of indivisibles is wrong?

Answer. What this example shows is that one cannot rely on intuition on the infinitesimal level, at least not until it has been chastened and corrected—"debugged," if you will. Logical clarity and rigor is the result of a debugging process. It leads to ever-increasing confidence that what we are doing is correct; but categorical, unqualified confidence is rash, even today. In the present case, one needs to show that the approximation is good in both *relative* and absolute terms. That is the secret of the infinitesimal methods of calculus: the approximation is good not only in the sense of being absolutely small, but in the sense of being small *even in comparison with the small objects doing the approximating* .

16.6. In the passage quoted from the *Analyst*, Berkeley asserts that the experience of the senses provides the only foundation for our imagination. From that premise he concludes that we can have no understanding of infinitesimals. Analyze whether the premise is true, and if so, whether it implies the conclusion. Assuming that our thinking processes have been shaped by the evolution of the brain, for example, is it possible that some of our spatial and counting intuition is "hard-wired" and not the result of any previous sense impressions? The philosopher Immanuel Kant (1724–1804) thought so. Do we have the power to make correct judgments about spaces and times on scales that we have not experienced? What would Berkeley have said if he had heard Riemann's argument that space may be finite, yet unbounded? How would he have explained the modern computer chip, on which unimaginable amounts of data can be recorded in space far too small for the senses to perceive? Go a step further and consider how quantum mechanics is understood and interpreted.

Answer. Berkeley makes a very good point here. No doubt our picture of the nucleus of an atom, for example, is not strictly speaking true. It is based on the use of delicate measuring instruments whose results we interpret in the context of physical theories. There is no sense in which a photograph of an atom made with a super-sensitive microscope is "what you would see if you were that small." In the first place if your size were comparable to that of an atom, you wouldn't have eyes. That point may seem naive, but it does show that the way we tend to think (at least the way I personally tend to think) about physical theories is not to be taken as a literal description of reality.

Kant really did think that the propositions of arithmetic and geometry were what he called "synthetic *a priori*"[1] judgments. That is, they were not mere logical tautologies—they were "synthetic"—obtained by joining two independent concepts to get a new one. His examples were the arithmetic relation $7 + 5 = 12$ and the geometric proposition that "a straight line is the shortest distance between two points." Here the union of a group of 7 and a group of 5 is one concept and a group of 12 is another concept, logically (so Kant believed) independent of the first. Thus the proposition "$7 + 5 = 12$" is synthetic rather than analytic. Yet, he said, it is a necessary result of the human thought process and is not derived from experience. Therefore it represents some knowledge about the world—universally true and necessarily true—that is innate in the human mind. Similarly in geometry, the notion of shortest distance and the notion of a straight line, Kant believed, were independent of each other, so that the proposition "a straight line is the shortest distance between two points" is synthetic knowledge. Again, our intuition rebels and will not accept the assertion that a straight line is *not* the shortest distance between two points, and therefore this knowledge is universal and necessary, hence *a priori*. Kant remarked that analytic propositions, such as syllogisms, do not increase human knowledge beyond their component parts. Thus, if I say "all children are young," I am not saying anything that was not already known to a person who knows what a child is and what it means to be young, since a child, by definition, is a young person. But synthetic propositions, he believed, really do add to human knowledge.

Now a modern mathematician and most modern mathematical philosophers would have little patience with all this, pointing out that the concepts involved in Kant's examples of synthetic *a priori* propositions have not been defined. When they are properly defined, as in modern geometry and set theory, they are seen to be analytic propositions. When we get a proper definition of distance, in terms of an integral, we can prove that the geodesics in Euclidean space are straight lines, and in set theory, when we define number properly, we can prove that $7 + 5 = 12$. In that way, these propositions are seen to be analytic,

[1] The Latin phrase *a priori* means *from earlier*, that is, *anterior* to any human experience and so not derived from observation or experience. Its opposite is *a posteriori* (*from later*), which is synonymous with *empirical*.

not synthetic. In Kant's defense, I would say that the definitions themselves conceal an unconscious assumption that these propositions are true, so that the propositions are in a sense circular reasoning.

For example, we define the length of a curve in Euclidean space as the *upper bound* of the lengths of all broken *straight lines* inscribed in it. That definition thus makes it trivial that the shortest distance between two points will be a straight line. But would we have made such a definition if we did not already have in mind that a straight line *is* the shortest distance? Similarly, you cannot tell anyone which symbols in set theory define the number 7 in a way that could be understood by someone who did not already know what the number 7 is. To be specific, we define the number 0 to be the empty set $\varnothing$, the number 1 to be the set $\{\varnothing\}$, defined so that $x \in 1$ if and only if x is a set and $x = 0$. We then define the number 2 to be $\{\varnothing, \{\varnothing\}\}$, that is, $x \in 2$ if and only if x is a set and $x = 0$ or $x = 1$. But unless a person already knows how to count to seven, it would be useless to try to explain that 7 is a set such that $x \in 7$ if and only if x is a set and $x = 0$ or $x = 1$ or $x = 2$ or $x = 3$ or $x = 4$ or $x = 5$ or $x = 6$. How would you know when you reached 7 if you didn't already know how to count? Actually the definition of 7 given here is a theorem. We should define 7 as the successor of 6. That is, $x \in 7$ if and only if x is a set and $x = 6$ or $x \in 6$. That definition would have meaning only if 6 were already defined as the successor of 5. So, doing our best for the set theorist/logicists (discussed in Chapter 19), you could tell someone what the number 7 is in this way, and they could trace it back to its source in the empty set by following along the route already laid out by the person who gave the definition. But the hypothetical person trying to find out what 7 is in this way doesn't know that the process ever terminates, or, even if it does, what information will be needed to get down to "bedrock." That information has to be supplied by the person communicating the definition. And it is obvious (is it not?) that this definition is a psychological monstrosity. Nobody really thinks about the number 7 this way. The chasm between the formal definition of 7 and the familiar number 7 is enormous.

I find attempts to explain what an ordered pair is similarly circular and fallacious. Some mathematics textbooks define the ordered pair (a, b) to be the set $\{a, \{a, b\}\}$, in which the element of the set that has no elements of its own is the first element and, by default, the other element is the second. But how would you know that (a, b) is not necessarily the same thing as (b, a)? Before you say that $(b, a) = \{b, \{b, a\}\}$, pause to reflect that you couldn't make this translation from the one symbol to the other unless you *already* knew that the order of the symbols on the page was part of the definition of (a, b).

I think Kant would have been unfazed by such attempts to make mathematics an analytic subject. He knew that geometry and arithmetic could be used to produce deductive systems that *model* the physical world. I think he would have regarded the modern mathematical approach as simply another application of such deductive models to the case of his synthetic *a priori* judgments—models of them, but not to be mistaken for them, any more than we think that a geometric line really *is* a light ray.

Against Kant's view[2] is the hard fact that human intuition is Euclidean, but geometry need not be. Kant would have taken Euclid's parallel postulate as a synthetic a priori proposition. Not that that would have ruled out the logical consistency of non-Euclidean geometry. Indeed, if non-Euclidean geometry were *logically* inconsistent, then Euclidean geometry would consist of analytic propositions. But non-Euclidean geometry is (in my view) fatal to Kant's philosophy even so, since one really can develop a non-Euclidean

[2] I would very much like to believe that Kant was right. His philosophy seems quite profound to me. But I reluctantly conclude that he was mistaken.

intuition and think in terms of hyperbolic space. This shows, at least to me, that we do not know the propositions of geometry innately. In addition, Kant held that ethical propositions are synthetic a priori judgments. He thought that the knowledge of right and wrong was innate in human beings and that it was factual knowledge. It seems to me far more likely that what is innate in human beings, as opposed to (say) orangutans, is a very strong tendency to seek social approval, so strong that, when social disapproval is supplemented by legal sanctions against violators, the overwhelming majority of people conform and only a few criminals defy the social code. That innate tendency will cause people to acquire the kind of subjective feeling of certainty about moral propositions that they acquire about arithmetic and geometry from everyday experience of counting and moving about. What the underlying brain mechanism is for arithmetic and geometry is a matter for psychologists and physiologists to ponder.

Modern technology and physics bring us increasingly face to face with the limitations of our intuition. The computer on which I am writing these words has a hard drive that stores 65 billion bytes of information. When I calculate the area of that hard drive, I find I really cannot imagine finding the location of the cell containing the code for the period at the end of this sentence. That code “looks like” 00101110 (binary code for the number 46) and I picture it as a row of eight tiny magnets, with north poles pointing upward in the first, second, fourth, and last places and downward in the other four. But the word *tiny* does not do justice to the smallness of those magnets. To get 65 billion sets of eight magnets on a disk, one must make them literally “unimaginably” small. We can’t imagine what they are “really” like, but we can imagine what they “would” look like (there we go again, into the subjunctive mood!) if suitably magnified. We can reason about them, and obviously our reasoning is good, since this computer works very well.

When we go to the still smaller world of quantum mechanics, we find even stranger things, things stranger than we even *can* imagine, as a physicist has said. All we know about quantum systems is their mathematics, yet that mathematics enables scientists to make very accurate predictions.

CHAPTER 17

Real and Complex Analysis

17.1. The familiar formula $\cos\theta = 4\cos^3(\theta/3) - 3\cos(\theta/3)$, can be rewritten as

$$p(\cos\theta/3, \cos\theta) = 0\,,$$

where $p(x, y) = 4x^3 - 3x - y$. Observe that $\cos(\theta + 2m\pi) = \cos\theta$ for all integers m, so that

$$p\left(\cos\left(\frac{\theta + 2m\pi}{3}\right), \cos\theta\right) \equiv 0\,,$$

for all integers m. That makes it very easy to construct the roots of the equation $p(x, \cos\theta) = 0$. They must be $\cos\big((\theta + 2m\pi)/3\big)$ for $m = 0, 1, 2$. What is the analogous equation for dividing a circular arc into five equal pieces?

Suppose (as is the case for elliptic integrals) that the inverse function of an integral is doubly periodic, so that $f(x + m\omega_1 + n\omega_2) = f(x)$ for all m and n. Suppose also that there is a polynomial $p(x)$ of degree n^2 such that $p\big(f(\theta/n)\big) = f(\theta)$. Show that the roots of the equation $p(x) = f(\theta)$ must be $f(\theta/n + (k/n)\omega_1 + (l/n)\omega_2)$, where k and l range independently from 0 to $n - 1$.

Answer. By the multi-angle formula $\cos(5\theta) = 16\cos^5\theta - 20\cos^3\theta + 5\cos\theta$. We take $p(x, y) = 16x^5 - 20x^3 + 5x - y$. Then the equation $p(x, \cos\theta) = 0$ has the five solutions $\cos\big((\theta + 2\pi k)/5\big)$, $k = 0, 1, 2, 3, 4$.

The second result is a consequence of the assumption that the equation holds for all θ. The right-hand side does not change when θ is replaced by $\theta + k\omega_1 + l\omega_2$, and hence the left-hand side must also be unchanged.

17.2. Show that if $y(x, t) = \big(f(x+ct) + f(x-ct)\big)/2$ is a solution of the one-dimensional wave equation that is valid for all x and t, and $y(0, t) = 0 = y(L, t)$ for all t, then $f(x)$ must be an odd function of period $2L$.

Answer. Taking $x = 0$, we find by the equation $y(0, t) \equiv 0$ that $f(ct) + f(-ct) \equiv 0$, which says that $f(x)$ is an odd function. Then, since $y(L, t) \equiv 0$, we find that $f(L + ct) + f(L - ct) = 0$. Taking $t = (x + L)/c$, we find that $0 = f(x + 2L) + f(-x) = f(x + 2L) - f(x)$ for all x, so that $f(x)$ has period $2L$.

17.3. Show that the problem $X''(x) - \lambda X(x) = 0$, $Y''(y) + \lambda Y(y) = 0$, with boundary conditions $Y(0) = Y(2\pi)$ and $Y'(0) = Y'(2\pi)$ has nonzero solutions only when $\lambda = n^2$, where n is an integer, and that the function $X(x)Y(y)$ must be of the form $\big(c_n e^{nx} + d_n e^{-nx}\big)\big(a_n\cos(ny) + b_n\sin(ny)\big)$ if $n \neq 0$.

Answer. It is impossible to get a nonzero solution satisfying these boundary condition if $\lambda < 0$. For let $\lambda = -\omega^2$. The equation $Y''(y) - \omega^2 Y(y) = 0$ has as solutions only functions of the form $Y(y) = ae^{\omega y} + be^{-\omega y}$. The conditions $Y(0) = Y(2\pi)$ and $Y'(0) = Y'(2\pi)$ then lead to a nondegenerate system of linear homogeneous equations for a and b, and so $a = b = 0$.

If $\lambda = 0$, the equation is $Y''(y) = 0$, which means that $Y(y) = ay + b$. The conditions $Y(0) = Y(2\pi)$ then imply that and $a = 0$. The solution in this case is of the form $X(x)Y(y) = cx + d$.

If $\lambda > 0$, then the solutions are of the form $X = ce^{\sqrt{\lambda}x} + de^{-\sqrt{\lambda}x}$, and $Y = a\cos(\sqrt{\lambda}y) + b\sin(\sqrt{\lambda}y)$. But the conditions $Y(0) = Y(2\pi)$ and $Y'(0) = Y'(2\pi)$ then imply $\lambda = n^2$ for some integer n, as shown in Problem 17.5 below.

17.4. Show that the differential equation

$$\frac{dx}{\sqrt{1-x^4}} + \frac{dy}{\sqrt{1-y^4}} = 0$$

has the solution $y = [(1-x^2)/(1+x^2)]^{1/2}$. Find another obvious solution of this equation.

Answer. The first assumption is mere computation:

$$\frac{dy}{dx} = \frac{1}{2}\frac{1}{y}\frac{d}{dx}\left(\frac{1-x^2}{1+x^2}\right) = \frac{1}{y}\frac{-2x}{(1+x^2)^2} = \frac{-2x}{(1+x^2)^{3/2}(1-x^2)^{1/2}}\,.$$

We need to show that this last expression is the same as

$$-\frac{\sqrt{1-y^4}}{\sqrt{1-x^4}}\,.$$

At least the negative sign is present in both expressions. Notice that

$$\sqrt{1-y^4} = \left(1 - \frac{(1-x^2)^2}{(1+x^2)^2}\right)^{1/2} = \left(\frac{4x^2}{(1+x^2)^2}\right)^{1/2} = \frac{2x}{1+x^2}\,.$$

Since $\sqrt{1-x^4} = \sqrt{(1-x^2)(1+x^2)}$, we find, as required, that

$$-\frac{\sqrt{1-y^4}}{\sqrt{1-x^4}} = \frac{-2x}{(1+x^2)^{3/2}(1-x^2)^{1/2}}\,.$$

The obvious solution is $y = -x$.

17.5. Show that Fourier series can be obtained as the solutions to a Sturm–Liouville problem on $[0, 2\pi]$ with $p(x) = r(x) \equiv 1$, $q(x) = 0$, with the boundary conditions $y(0) = y(2\pi)$, $y'(0) = y'(2\pi)$. What are the possible values of λ?

Answer. These values of p, q, and r, give the equation $y''(t) + \lambda y(t) = 0$, whose solutions for positive λ are $y(t) = a\cos(\sqrt{\lambda}t) + b(\sin\sqrt{\lambda}t)$. The boundary conditions $y(0) = y(2\pi)$, $y'(0) = y'(2\pi)$ lead to a system of linear equations for a and b, which will have only trivial solutions unless the coefficient matrix is singular. The determinant of that matrix is $2\sqrt{\lambda}\bigl(1-\cos(2\pi\sqrt{\lambda})\bigr)$. Thus $\cos(2\pi\sqrt{\lambda})$ must equal 1, and so $2\pi\sqrt{\lambda} = 2\pi n$ for some integer n. Thus $\sqrt{\lambda} = n$. Obviously, the solutions are constants if $\lambda = 0$, and, as shown in Problem 17.3 above, there are no nonzero solutions if $\lambda < 0$.

CHAPTER 18

Probability and Statistics

18.1. Weather forecasters are evaluated for accuracy using the *Briers score.* The *a posteriori* probability of rain on a given day, judged from the observation of that day, is 0 if rain did not fall and 1 if rain did fall. A weather forecaster who said (the day before) that the chance of rain was 30% gets a Briers score of $30^2 = 900$ if no rain fell and $70^2 = 4900$ if rain fell. Imagine a very good forecaster, who over many years of observation learns that a certain weather pattern will bring rain 30% of the time. Also assume that for the sake of negotiating a contract that forecaster wishes to optimize (minimize) his or her Briers score. Should that forecaster state truthfully that the probability of rain is 30%? If we assume that the prediction and the outcome are independent events, we find that, for the days on which the true probability of rain is 30% the forecaster who makes a prediction of a 30% probability would in the long run average a Briers score of $0.3 \cdot 70^2 + 0.7 \cdot 30^2 = 2100$. This score is better (in the sense of a golf score—it is lower) than would result from randomly predicting a 100% probability of rain 30% of the time and a 0% probability 70% of the time. That strategy will be correct an expected 58% of the time ($.58 = .3^2 + .7^2$) and incorrect 42% of the time, resulting in a Briers score of $.42 \cdot 100^2 = 4200$. Let p be the actual probability of rain and x the forecast probability. Assuming that the event and the forecast are independent, show that the expected Briers score $10^4\big(p(1-x)^2 + (1-p)x^2\big)$ is minimized when $x = p$. [*Note*: If this result did not hold, a meteorologist who prized his/her reputation as a forecaster, based on the Briers measure, would be well advised to predict an incorrect probability, so as to get a better score for accuracy!]

Answer. The Briers score is $10^4(1-x)^2$ on a day when rain fell when the forecaster gave a probability of x. It is 10^4x^2 if no rain fell. Hence, over a period of N days when the actual probability of rain is p and the predicted probability is x ("actual" meaning that it really will rain on pN of these days) the forecaster will accumulate an overall Briers score of $10^4N\big(p(1-x)^2 + (1-p)x\big)$, for an average Briers score per day of $10^4\big(p(1-x)^2 + (1-p)x^2\big)$. This is a quadratic function of x, equal to

$$10^4\big(x^2 - 2px + p\big) = 10^4\big((x-p)^2 + p(1-p)\big)\,,$$

whose unique minimum value $10^4p(1-p)$ occurs when $x = p$.

18.2. We saw above that Cardano (probably) and Pascal and Leibniz (certainly) miscalculated some elementary probabilities. As an illustration of the counterintuitive nature of many simple probabilities, consider the following hypothetical games. (A casino could probably be persuaded to open such games if there was enough public interest in them.) In game 1 the dealer lays down two randomly-chosen cards from a deck on the table and turns one face up. If that card is not an ace, no game is played. The cards are replaced in the deck, the deck is shuffled, and the game begins again. If the card is an ace, players are invited to bet against a fixed winning amount offered by the house that the other card is also an ace. What winning should the house offer (in order to break even in the long run) if players pay one dollar per bet?

In game 2 the rules are the same, except that the game is played only when the card turned up is the ace of hearts. What winning should the house offer in order to break even charging one dollar to bet? Why is this amount not the same as for game 1?

Answer. For the first game, most of the hands dealt will not be played. There are only 198 different equally likely and playable hands, of which the house will win 192. Hence the odds are 32 to 1 against the players. The house could offer $32 to the winning gambler and lose only its overhead by doing so.

For the second game, even fewer of the hands dealt will be played, 51 to be exact. Most of the games that would have been played previously but will not be played under the new rules are games that the players would have lost. Under the new rules the house will win 48 of the games and the players 3, so the odds are 16 to 1 against the players, twice as good as in the other case.

18.3. Use the Maclaurin series for $e^{-(1/2)t^2}$ to verify that the series given by de Moivre represents the integral

$$\frac{1}{\sqrt{2\pi}}\int_0^1 e^{-\frac{1}{2}t^2}\,dt,$$

which is the area under a standard normal (bell-shaped) curve within one standard deviation of the mean, as given in many tables.

Answer. The Maclaurin series is

$$e^{-(1/2)t^2} = 1 - \frac{t^2}{2} + \frac{t^4}{2^4 2!} - \frac{t^6}{2^6 3!} + \frac{t^8}{2^8 4!} - \frac{t^{10}}{2^{10} 5!} + \cdots .$$

Hence the integral is

$$\frac{1}{\sqrt{2\pi}}\int_0^1 e^{-(1/2)t^2}\,dt = \frac{1}{\sqrt{2\pi}}\left(1 - \frac{1}{2\cdot 3} + \frac{1}{5\cdot 2^4\cdot 2!} - \frac{1}{7\cdot 2^6 3!} + \cdots\right).$$

However, I ought to have displayed the application of this series given by de Moivre, who said that the sum of the terms in the expansion of $\left(\frac{1}{2}+\frac{1}{2}\right)^{2n}$ having binomial coefficients between $\binom{2n}{n}$ and $\binom{2n}{n+l}$ would be (approximately)

$$\frac{2}{\sqrt{nc}}\left(l - \frac{2l^3}{1!\cdot 3n} + \frac{4l^5}{2!\cdot 5n^2} - \frac{8l^7}{3!\cdot 7n^4} + \cdots\right),$$

where c is "the circumference of a circle whose radius is unity," that is, $c = 2\pi$. He then took $l = s\sqrt{n}$, getting the expression we would write as

$$\sqrt{\frac{2}{\pi}}\left(s - \frac{2s^3}{1!\cdot 3} + \frac{4s^5}{2!\cdot 5} - \frac{8s^7}{3!\cdot 7} + \cdots\right),$$

Finally, he set $s = \frac{1}{2}$, getting

$$\sqrt{\frac{2}{\pi}}\left(\frac{1}{2} - \frac{1}{1!\cdot 3\cdot 4} + \frac{1}{2!\cdot 5\cdot 8} - \frac{1}{3!\cdot 7\cdot 16} + \cdots\right),$$

as the sum of the terms between the middle and $\frac{1}{2}\sqrt{n}$ from the middle. This is trivially the same as the expression given above for the integral.

18.4. Use Daniel Bernoulli's concept of utility to explain why only a person with astronomical amounts of money should play a Petersburg paradox-type game. In your explanation, take account of what the utility of the stakes must be for a gambler versus the utility of the gain. Make an analogy between risk and work in this regard. A laborer exchanges time and effort for money; a gambler exchanges risk for potential gain. Remembering that all economic decisions are made "at the margin," at what point does additional work (or risk) not bring enough additional utility to be worth the exchange?

Answer. Whether or not Daniel Bernoulli's specific formula for utility is valid, the phenomenon of diminishing returns or decreased marginal utility is well established qualitatively. For a gambler with a small stake, the marginal utility of the money to be put at risk is very high. What Bernoulli's formula recognizes is that the marginal utility of a large but improbable gain is not high. A strict adherence to mathematical expectation may account for the purely monetary aspects of the game. But, just as the apparent size of distant objects is not inversely proportional to their distance (as discussed in Chapter 10), the utility (analogous to apparent size) of a win is not directly proportional to its probability. It decreases faster than that as the probability decreases.

18.5. Radium-228 is an unstable isotope. Each atom of Ra-228 has a probability of 0.1145 (about 1 chance in 9, or about the probability of rolling a 5 with two dice) of decaying to form an atom of actinium within any given year. This means that the probability that the atom will survive the year as an atom of Ra-228 is $1 - 0.1145 = 0.8855$. Denote this "one-year survival" probability by p. Because any sample of reasonable size contains a huge number of atoms, that survival probability (0.8855) is the proportion of the weight of Ra-228 that we would expect to survive a year.

If you had one gram of Ra-228 to begin with, after one year you would expect to have $p = 0.8855$ grams. Each succeeding year, the weight of the Ra-228 left would be multiplied by p, so that after two years you would expect to have $p^2 = (0.8855)^2 = 0.7841$ grams. In general, after t years, if you started with W_0 grams, you would expect to have $W = W_0 p^t$ grams. Now push these considerations a little further and determine *how strongly* you can rely on this expectation. Recall Chebyshëv's inequality, which says that the probability of being more than k standard deviations from the expected value is never larger than $(1/k)^2$. What we need to know to answer the question in this case is the standard deviation σ.

Our assumption is that each atom decays at random, independently of what happens to any other atom. This independence allows us to think that observing our sample for a year amounts to a large number of "independent trials," one for each atom. We test each atom to see if it survived as an Ra-228 atom or decayed into actinium. Let N_0 be the number of atoms that we started with. Assuming that we started with 1 gram of Ra-228, there will be $N_0 = 2.642 \cdot 10^{21}$ atoms of Ra-228 in the original sample.[1] That is a very large number of atoms. The survival probability is $p = 0.8855$. For this kind of independent trial, as mentioned the standard deviation with N_0 trials is

$$\sqrt{N_0 p(1-p)} = \sqrt{\frac{p(1-p)}{N_0}} N_0 \,.$$

[1] According to chemistry, the number of atoms in one gram of Ra-228 is the *Avogadro number* $6.023 \cdot 10^{23}$ divided by 228.

We write the standard deviation in this odd-looking way so that we can express it as a fraction of the number N_0 that we started with. Since weights are proportional to the number of atoms, that same fraction will apply to the weights as well.

Put in the given values of p and N_0 to compute the fraction of the initial sample that constitutes one standard deviation. Since the original sample was assumed to be one gram, you can regard the answer as being expressed in grams. The use Chebyshëv's inequality to estimate the probability that the amount of the sample remaining will differ from the theoretically predicted amount by 1 millionth of a gram (1 microgram, that is, 10^{-6} grams)? [*Hint*: How many standard deviations is one millionth of a gram?]

Answer. My calculator gives the standard deviation as $6.1948 \times 10^{-12} N_0$ atoms, which in our case amounts to 6.1948×10^{-12} grams. Hence the hypothetical deviation of 10^{-6} grams amounts to over 100,000 standard deviations. According to Chebyshëv's inequality, the probability of being this far from the expected value is less than 10^{-10}. That is, the odds that you will observe a deviation this large are far less than 1 in ten billion.

18.6. Analyze the revised probabilities in the problem of two drawers, one containing two gold coins, the other a gold and a silver coin, given an experiment in which event B occurs, if B is the event, "a silver coin is drawn."

Answer. Since there is no probability that you are in the drawer containing two gold coins, the probabilities should now be 0 and 1. And indeed they are. Let A be the event, "The drawer contains two gold coins," and C the event, "The drawer contains a gold coin and a silver coin." Then

$$\begin{aligned} P(A \big| B) &= P(A \cap B)/P(B) = 0/0.25 = 0\,; \\ P(C \big| B) &= P(C \cap B)/P(B) = 0.25/0.25 = 1\,. \end{aligned}$$

18.7. Consider the case of 200 men and 200 women applying to a university consisting of only two different departments, and assume that the acceptance rates are given by the following table.

	Men	**Women**
Department A	120/160	32/40
Department B	8/40	40/160

Observe that the admission rate for men in department A is $\frac{3}{4}$, while that for women is $\frac{4}{5}$. In department B the admission rate for men is $\frac{1}{5}$ and for women it is $\frac{1}{4}$. In both cases, the people actually making the decisions are admitting a higher proportion of women than of men. Now explain the source of the bias, in our example and at Berkeley in simple, nonmathematical language.

Answer. Women as a group are applying to departments that are more difficult to get into. The selection process at the university is not necessarily biased, but the applicant group itself is biased in its selection of departments.

CHAPTER 19

Logic and Set Theory

19.1. Bertrand Russell pointed out that some applications of the axiom of choice are easier to avoid than others. For instance, given an infinite collection of pairs of shoes, describe a way of choosing one shoe from each pair. Could you do the same for an infinite set of pairs of socks?

Answer. As you can readily see, the problem is one of distinguishing individual elements of a set. If each element of a set is in some way distinguishable from every other one, then one can *order* the set. In the case of infinitely many pairs of shoes, since left and right shoes are (normally) distinguishable, one can put "left before right" and then choose the left shoe from each pair. In general, if sets are well-ordered, one can choose the first element of each set. However, as we saw, the assumption that every set can be well-ordered is equivalent to the axiom of choice.

19.2. Prove that $C = \{x : x \notin x\}$ is a proper class, not a set, that is, it is not an element of any class.

Answer. The assumption that C is an element of itself would imply (by its definition) that it is a set and is *not* an element of itself, which is a contradiction. If it is a set and is not an element of itself, then by its own definition, it *is* an element of itself, again a contradiction and a paradox. But if we grant that C is a proper class, we avoid the contradiction (as already stated in the text) since the assumption that it is not an element of itself is not in contradiction with its definition.

19.3. Suppose that the only allowable way of forming new formulas from old ones is to connect them by an implication sign; that is, given that A and B are well formed, $[A \Rightarrow B]$ is well formed, and conversely, if A and B are not both well formed, then neither is $[A \Rightarrow B]$. Suppose also that the only basic well-formed formulas are p, q, and r. Show that

$$\Big[[p \Rightarrow r] \Rightarrow \big[[p \Rightarrow q] \Rightarrow r\big]\Big]$$

is well formed but

$$\big[[p \Rightarrow r] \Rightarrow [r \Rightarrow]\big]$$

is not. Describe a general algorithm for determining whether a finite sequence of symbols is well formed.

Answer. By the rules for formation of well-formed formulas, $[p \Rightarrow r]$ is well formed. The formula $[p \Rightarrow q]$ is also well formed, and hence so is $\big[[p \Rightarrow q] \Rightarrow r\big]$. Thus finally the formula

$$\Big[[p \Rightarrow r] \Rightarrow \big[[p \Rightarrow q] \Rightarrow r\big]\Big]$$

is also well formed.

As a general algorithm, we note that a formula is a finite sequence of symbols chosen from the set $\{[,], p, q, r, \Rightarrow\}$. The first test in the algorithm is the following: *If a formula contains no occurrences of $\Rightarrow$, it must be a single symbol chosen from $\{p, q, r\}$. If the formula is one of these three symbols, the algorithm terminates.*

If the formula contains the symbol $\Rightarrow$, it passes the first test, but the algorithm proceeds to the second test: *The formula must begin with the left bracket and end with the right bracket, and the number of left brackets must equal the number of right brackets.*

If a formula passes the second test, we apply a recursion that reduces the number of occurrences of the symbol $\Rightarrow$: *All but one of the occurrences of the symbol $\Rightarrow$ must be between two or more pairs of left and right brackets, that is, exactly one occurrence of $\Rightarrow$ must be outside all pairs of brackets except the outer pair.* If the formula fails this test, the algorithm terminates. If it passes, the algorithm moves recursively to two sequences of symbols: the sequence preceding the occurrence of $\Rightarrow$ just discussed and the sequence following it. The original formula is well formed if and only if both of these sequences of symbols are well formed. Since each of these formulas has fewer occurrences of $\Rightarrow$ than the original formula, this algorithm must terminate in a finite number of steps. If at any stage a formula is reached that is not well formed, the original formula is not well formed; otherwise it is.

19.4. Consider the following theorem. There exists an irrational number that becomes rational when raised to an irrational power. *Proof:* Consider the number $\theta = \sqrt{3}^{\sqrt{2}}$. If this number is rational, we have an example of such a number. If it is irrational, the equation $\theta^{\sqrt{2}} = \sqrt{3}^2 = 3$ provides an example of such a number. Is this proof intuitionistically valid?

Answer. The proof is not intuitionistically valid, since it asserts that "p or q" is true without proving either that p is true or that q is true. In the intuitionistic propositional calculus, if $p \vee q$ is a theorem, then either p is a theorem or q is a theorem.

19.5. Show that any two distinct *Fermat numbers* $2^{2^m} + 1$ and $2^{2^n} + 1$, $m < n$, are relatively prime. (Use mathematical induction on n.) Apply this result to deduce that there are infinitely many primes. Would this proof of the infinitude of the primes be considered valid by an intuitionist?

Answer. Consider the numbers $G_m = F_m - 1 = 2^{2^m}$ and observe that $G_{m+1} = G_m^2$. Hence $G_{m+k} = G_m^{2^k}$. This equality asserts that

$$F_{m+k} = 1 + (F_m - 1)^{2^k} ,$$

from which it follows by the binomial theorem that

$$F_{m+k} = F_m^{2^k} - 2^k F_m^{2^k-1} + \cdots - 2^k F_m + 2 = QF_m + 2 ,$$

That is, each Fermat number, when divided by a smaller Fermat number, leaves a remainder of 2. Thus the only possible common divisors of two Fermat numbers are 1 and 2. Since all Fermat numbers are odd, 2 cannot be a common divisor.

Hence if we take all the prime divisors of Fermat numbers, we must obtain an infinite set of primes. This is short of exhibiting an algebraic formula that always generates a prime,[1] but it does give an *algorithm* in which each iteration produces a new prime. The algorithm proceeds as follows. Form the number F_m. Then divide F_m by each positive

[1] Such a formula (an algebraic polynomial with integer coefficients) was constructed in principle by Matiyasevich in the course of his solution of Hilbert's Tenth Problem, as discussed in the answer to Question 4.3.

integer, starting with 3, until the first integer is reached at which the remainder is zero. Let that integer, which is necessarily prime and necessarily less than or equal to F_m, be p_m. Increment m and continue.

This algorithm ought to satisfy an intuitionist, who should confess that the primes are at least potentially infinite.

19.6. Suppose that you prove a theorem by assuming that it is false and deriving a contradiction. What you have then proved is that either the axioms you started with are inconsistent or the assumption that the theorem is false is itself false. Why should you conclude the latter rather than the former? Is this why some mathematicians have claimed that the practice of mathematics requires faith?

Answer. The words *faith* and *should* are slippery ones. The agnostic position is always available to both scientists and mathematicians: It is possible to *explore the consequences* of a proposition without *affirming* the proposition. This position is not available in other areas, and it contradicts the meaning of the word *faith*. Mathematicians who use set theory, for example, can state that they are using it only to derive theorems and make no claim as to its consistency. In that respect a mathematician need not assert that we *should* draw any conclusions at all from our proofs, other than the hypothetical conclusion that "if all our assumptions are true (and hence consistent with one another), then our conclusions are also true." Thus it can be argued that the word *faith* is misapplied in both mathematics and science, at least as far as pure logic is concerned.

Where logic is satisfied, however, human psychology is not. If mathematicians did not have considerable *confidence*[2] in the consistency of set theory, they would not use it, any more than chemists and physicists would devote large amounts of time and effort seeking a reaction (cold fusion, for example) that they did not believe possible. Thus the word *faith* comes back on the psychological level. It is a rather anemic faith, however, compared with the degree of conviction that some religions expect their followers to have in reponse to a detailed creed. Confidence exists in various degrees, expressed as probabilities: One can bet at odds of arbitrarily high levels on the correctness of the multiplication table. I personally would want much better odds before I would endorse every page in the classification of the finite groups.

19.7. What are the advantages, if any, of building a theory by starting with abstract definitions, then later proving a structure theorem showing that the abstract objects so defined are actually familiar objects?

Answer. The chief advantage is that the abstract approach reveals the logical structures by which the theory produces its results. As long as mathematicians were confined to dividing integers and polynomials, for example, they were hindered in seeing the underlying "valuations"—$V(n) = |n|$ for an integer n, $V(p) = 2^n$, where n is the degree of the polynomial (by convention $V(0) = 0$)—that make the Euclidean algorithm applicable. Another advantage is that one sees the contexts in which an argument can work and can avoid trying to apply it where it won't work. On that basis, it is worthwhile studying finite-dimensional division algebras over the real numbers to see what properties they have, even though the only possible examples are the real numbers themselves, the complex numbers, and the quaternions.

[2] This is a "faith-based" word. It comes from the Latin word *fides*, meaning *faith*.

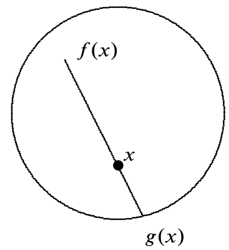


FIGURE 1. The Brouwer fixed-point theorem.

19.8. Brouwer, the leader of the intuitionist school of mathematicians, is also known for major theorems in topology, including the invariance of geometric dimension under homeomorphisms and the *Brouwer fixed-point theorem*, which asserts that for any continuous mapping f of a closed disk into itself there is a point x such that $x = f(x)$. To prove this theorem, suppose there is a continuous mapping f for which $f(x) \neq x$ at every point x. Construct a continuous mapping g by drawing a line from $f(x)$ to x and extending it to the point $g(x)$ at which it meets the boundary circle (see Fig. 1). Then $g(x)$ maps the disk continuously onto its boundary circle and leaves each point of the boundary circle fixed. Such a continuous mapping is intuitively impossible (imagine stretching the entire head of a drum onto the rim without moving any point already on the rim and without tearing the head) and can be shown rigorously to be impossible (the disk and the circle have different homotopy groups). How can you explain the fact that the champion of intuitionism produced theorems that are not intuitionistically valid?

Answer. As mentioned in the text, Brouwer wasn't an intuitionist at the time when he produced this theorem. There are, as far as I know, no proofs of this fixed-point theorem that are valid on intuitionist principles, although I have seen one that is constructive except for requiring the reader to recognize whether a set is finite or infinite at certain points. Even for an intuitionist, however, these theorems proved using more general principles of inference can serve a purpose, suggesting what an intuitionist wouldn't try to do and what it might be interesting to try.

19.9. A naive use of the formula for the sum of the geometric series $1/(1 + x) = 1 - x + x^2 - x^3 + \cdots$ seems to imply that $1 - 5 + 25 - 125 + \cdots = 1/(1 + 5) = 1/6$. Nineteenth-century analysts rejected this use of infinite series and confined themselves to series that converge in the ordinary sense. However, Kurt Hensel (1861–1941) showed in 1905 that it is possible to define a notion of distance (the p-adic metric) by saying that an integer is close to zero if it is divisible by a large power of the prime number p (in the present case, $p = 5$). Specifically, the distance from m to 0 is given by $d(m, 0) = 5^{-k}$, where 5^k divides m but 5^{k+1} does not divide m. The distance between 0 and the rational number $r = m/n$ is then by definition $d(m, 0)/d(n, 0)$. Show that $d(1, 0) = 1$. If the distance between two rational numbers r and s is defined to be $d(r - s, 0)$, then in fact the series just mentioned does converge to $\frac{1}{6}$ in the sense that $d(S_n, \frac{1}{6}) \to 0$, where S_n is the nth partial sum.

What does this historical experience tell you about the truth or falsity of mathematical statements? Is there an "understood context" for every mathematical statement that can never be fully exhibited, so that certain assertions will be *verbally* true in some contexts and verbally false in others, depending on the meaning attached to the terms?

Answer. One way of getting around counterexamples is always to reinterpret the meaning of the terms. Thus, Cauchy had given a definition of continuity that looked reasonable and had given an argument that a limit of continuous functions is continuous. But he hadn't been as specific as he might have been when giving the definition of convergence, and Abel pointed out that there were cases where a sequence of continuous functions could converge to a discontinuous function *in one interpretation of the meaning of convergence*. It then took some time for different varieties of convergence to be identified. Once they were identified, Cauchy's theorem could be reinstated, but the word *convergence* would henceforth be replaced by the phrase *uniform convergence*.

In the present example, it is easy to verify that the difference between $\frac{1}{6}$ and the partial sum of the series is $\pm\frac{5^{n+1}}{6}$, whose distance from 0 is 5^{-n-1} and hence tends to zero.

19.10. Are there true but unknowable propositions in everyday life? Suppose that your class meets on Monday, Wednesday, and Friday. Suppose also that your instructor announces one Friday afternoon that you will be given a surprise exam at one of the regular class meetings the following week. One of the brighter students then reasons as follows. The exam will not be given on Friday, since if it were, having been told that it would be one of the three days, and not having had it on Monday or Wednesday, we would know on Thursday that it was to be given on Friday, and so it wouldn't be a surprise. Therefore it will be given on Monday or Wednesday. But then, since we *know* that it can't be given on Friday, it also can't be given on Wednesday. For if it were, we would know on Tuesday that it was to be given on Wednesday, and again it wouldn't be a surprise. Therefore it must be given on Monday, we know that now, and therefore it isn't a surprise. Hence it is impossible to give a surprise examination next week.

Obviously something is wrong with the student's reasoning, since the instructor can certainly give a surprise exam. Most students, when trying to explain what is wrong with the reasoning, are willing to accept the first step. That is, they grant that it is impossible to give a *surprise* exam on the *last* day of an assigned window of days. Yet they balk at drawing the conclusion that this argument implies that the originally next-to-last day must thereby become the last day. Notice that, if the professor had said nothing to the students, it would be possible to give a surprise exam on the last day of the window, since the students would have no way of knowing that there was any such window. The conclusion that the exam cannot be given on Friday therefore does not follow from assuming a surprise exam within a limited window alone, but rather from these assumptions supplemented by the following proposition: *The students* know *that the exam is to be a surprise and they* know *the window in which it is to be given.*

This fact is apparent if you examine the student's reasoning, which is full of statements about what the students *would know*. Can they truly *know* a statement (even a true statement) if it leads them to a contradiction?

Explain the paradox in your own words, deciding whether the exam would be a surprise if given on Friday. Can the paradox be avoided by saying that the conditions under which the exam is promised are true but the students cannot *know* that they are true?

How does this puzzle relate to Gödel's incompleteness result?

Answer. Whenever a system of propositions is capable of "talking about itself" and can meaningfully formulate the assertion that certain propositions can be proved within the system or cannot be proved within the system, the possibility of an incompleteness result of this type occurs. The crux of the matter is the meaning of the phrase *to know*. In informal speech we trust others enough to say that we know things that we have been told. But if someone tells us something that leads to a contradiction, we cease to include their word as a basis for knowing. In the present instance the students are tantalized by being told something by a teacher—ordinarily a sufficiently rigorous foundation for knowledge—that apparently leads to a contradiction *if they know it*. Hence they are in the situation of "P implies not-P". Here P is the proposition "we know that the two statements made by the instructor are true." If P is true, it leads to correct inferences by the students that contradict P. Therefore, any logician would agree, P is false. Notice that what is false is not what the students were told; that was always true and remains true. What is false is that the students know what they were told.

Gödel's incompleteness result is of exactly this form. It is a formula whose interpretation says that it cannot be proved. (Since in arithmetic proof and knowing are twins, the formula really says, "I cannot be known to be true.")

Subject Index

Name Index