

CALCULATING $\sqrt{2}$

Yutaka Nishiyama

Department of Business Information

Faculty of Information Management

Osaka University of Economics

2, Osumi Higashiyodogawa Osaka, 533-8533, JAPAN

Abstract: This article concerns root 2. The ratio of height and width of copy paper is root 2, and can be calculated by using similar diagrams. Root 2 is an irrational number whose cardinality is uncountably infinite. Roots are calculated numerically by the Newton-Raphson method which uses an approximating 2nd order function and its tangent.

AMS Subject Classification: 68W02, 00A09, 97A20

Key Words: ratio of height and width of copy paper, similar diagrams, irrational number, *Reductio ad absurdum*, cardinality of numbers, Cantor's diagonal method, countably infinite and uncountably infinite, 2nd order function and its tangent, Newton-Raphson method

1. The Height and Width of Photocopier Paper

I wondered exactly how much students in the humanities know about root 2, so I took this up in a lecture on information mathematics. Root 2 is written $\sqrt{2}$, and produces 2 when it is squared. I had my students discuss this value themselves to see what they would say. Students who didn't know its value already made the following kinds of calculation

$$1^2 = 1 \text{ and } 2^2 = 4 \text{ so } 1 < \sqrt{2} < 2$$

$$1.5^2 = 2.25 \text{ and } 1.4^2 = 1.96 \text{ so } 1.4 < \sqrt{2} < 1.5$$

They explained how to obtain the value of root 2 within a certain level of accuracy by means of written calculations using this bisection method. Some students knew the values of root 2 and root 3 without the need for calculation according to mnemonics like “I wish I knew -the root of two. O charmed was he - to know root three. So we now strive - to find root five” (which encodes the values 1.414, 1.732 and 2.235).

A4 and B4 copy paper sizes are of a familiar size, and the ratio of their height and width is 1 to root 2, although surprisingly few people know this. This ought to have been learned in junior high-school geometry classes or the first year of high school, but whether mathematics was not necessary for their examinations, or whether they just hated mathematics, many students forget this fact. I therefore give students the following problem to think about.

Problem 1. Find the ratio of the height and width of a piece of copy paper.

Students reply with ‘1 to 2’, or ‘2 to 3’ according to the appearance of the paper. At that point I demonstrate that copy paper has the amazing property that folding it in half does not change the ratio of the height and width. Such shapes are known as similar diagrams. Maybe students learn this for the first time, - they always seem newly impressed.

The ratio of the height and width may be obtained using this property of similarity. Let’s denote the ratio of the height and width of a rectangle as x to 1. Thinking about the same rectangle folded in half reveals that the height is $\frac{x}{2}$ while the width is 1, so

$$1 : x = \frac{x}{2} : 1.$$

Since the product of the inner terms is equivalent to the product of the outer terms (this should have been learned at junior high-school),

$$x \times \frac{x}{2} = 1 \times 1.$$

Solving this yields $x = \sqrt{2}$. That is, the ratio of the height and width of copy paper is $1 : \sqrt{2}$.

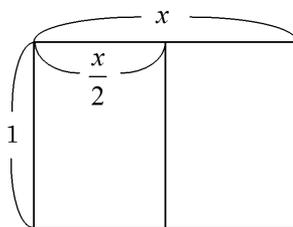


Figure 1: The ratio of the height and width of copy paper

2. Proving that Root 2 is an Irrational Number

It is known that root 2 is an irrational number. In response to this information, students ask what irrational numbers are, and I reply that they are numbers which continue forever after the decimal point. “Doesn’t that mean that numbers like $0.999999\dots$, which go on forever are irrational?” they ask me.

$$\frac{1}{3} = 0.333333\dots \quad \frac{1}{3} \times 3 = 0.999999\dots$$

so,

$$1 = 0.999999\dots$$

This is not an irrational number. To put it precisely, irrational numbers are those numbers which continue indefinitely after the decimal point, without cycling. Numbers like

$$\frac{2}{15} = 0.133333\dots, \text{ and } \frac{1}{7} = 0.1428571428571\dots$$

and so on, continue indefinitely but when the part immediately after the decimal point cycles, the number is not known as irrational. Recurring decimals can be expressed as vulgar fractions by calculating their value as an infinite series.

Rational numbers can be expressed as a ratio of integers, but irrational numbers cannot be expressed as an integer ratio. Few students can reply correctly up to this level. In Japanese, ‘irrational numbers’ are known as *murisu*, but the part corresponding to ‘irrational’ also has other meanings, and the name is not so appropriate. During the *Meiji* era, the English terms ‘rational number’ and ‘irrational number’ were translated into Japanese as *yurisu* and *murisu*, but the real meaning is whether or not they can be expressed as a ratio of integers, so some mathematical historians have suggested that the terms *yuhisu* and *muhisu* meaning respectively, ‘number with a fraction’ and ‘number without a fraction’, are more appropriate.

Problem 2. Prove that $\sqrt{2}$ is an irrational number.

Well then, let's try and prove that root 2 is irrational, *i.e.*, that it cannot be expressed as a ratio of integers. This is explained in the high-school mathematics syllabus I-A, but there is little chance of the proof appearing in university entrance examinations, so many students avoid studying it and do not know the method. It is important, so I'd like to reiterate the proof here. The proof utilizes *reductio ad absurdum*.

Suppose that root 2 could be expressed as a ratio of integers as follows.

$$\sqrt{2} = \frac{q}{p} \quad (\text{where } p \text{ and } q \text{ are mutually prime natural numbers})$$

The fact that p and q are mutually prime expresses a requirement that they constitute an irreducible fraction. Squaring both sides of this equation and eliminating the denominator yields

$$q^2 = 2p^2.$$

From this equation, q^2 is a factor of 2. If q^2 is a factor of 2, then q is also a factor of 2 (this part of the proof is a little tricky). We can thus write,

$$q = 2m \quad (\text{where } m \text{ is a natural number}).$$

Substituting this into the previous equation and rearranging yields the following.

$$4m^2 = 2p^2$$

$$p^2 = 2m^2$$

From this equation p^2 is a factor of 2. Again, if p^2 is a factor of 2 then p is also a factor of 2. Putting these results together, q is a factor of 2, and p is a factor of 2. This contradicts the assumption that p and q are mutually prime. Thus $\sqrt{2}$ cannot be expressed as a fraction $\frac{q}{p}$ (an integer ratio), and is therefore an irrational number.

3. The Cardinality of Rational Numbers and Irrational Numbers

How many rational numbers and how many irrational numbers are there? Which type describes the most numbers? At this point allow me to introduce a rather interesting topic. It should have been learned that according to Euclidean geometry “the sum of two edges of a triangle is larger than the other edge, and the difference between two edges is less than the other edge”. However, let's try to prove that “the sum of two edges of a triangle is equal to

the other edge” or alternatively “all the line segments have equal length?!” Of course this is sophistry, but can we identify where the mistake lies?

Suppose we have the triangle ABC shown in Figure 2. Draw $B'C'$ parallel to the base edge BC. In general “lines are a collection of points”, “faces are a collection of straight lines” and “solids are a collection of faces”. The line segment BC is packed with points. As a representative example let us select the point P. Connecting this point P and the vertex A produces a line that crosses $B'C'$, and the intersection point is denoted Q. Every point on line segment BC has a corresponding point on line segment $B'C'$. We postulated that a line segment was a collection of points, and thus the line segments BC and $B'C'$ are equivalent.

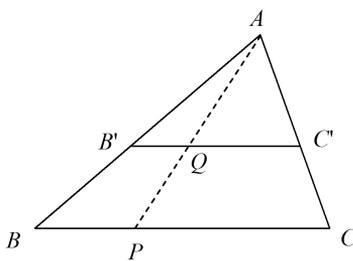


Figure 2: Are line segments BC and $B'C'$ equivalent?

Around the time I was a first year university student there was an aspect of mathematics that I had just learned, and all I wanted to do was talk about it with my friends. The particular piece of sophistry in question identified a mistake in the notion that “lines are a collection of points”. No matter how many points are gathered, they do not constitute a line. They are just a collection of points. This is just how difficult it is to comprehend infinity. In this regard, there are an infinite number of natural numbers, rational numbers and irrational numbers, but the natural numbers and rational numbers form infinite sets that may be enumerated (they are ‘countable’), while the irrational numbers form an infinite set that cannot be enumerated (they are ‘uncountable’).

In 1891, Cantor proved that there is no 1-1 mapping from the natural numbers to the real numbers in the interval $(0, 1]$ using the so-called diagonal method. This is because the cardinality of any set which has a 1-1 mapping with the set of natural numbers must be equal to the cardinality of the natural numbers. This cardinality is written \aleph_0 , and read ‘aleph zero’. There is a 1-1 correspondence between the rational numbers and the natural numbers, so the cardinality of the rational numbers is \aleph_0 . \aleph is a Hebrew letter corresponding

to the letter A. Sets with a cardinality of \aleph_0 are known as enumerable, or ‘countably infinite’. In contrast, the cardinality of the real numbers is written \aleph , and $\aleph_0 < \aleph$.

4. Tangential Line Equations

The explanation is now for the most part complete. The theme in this chapter was the calculation of root 2. If the 2nd order function and the corresponding tangent are known, then root 2 can be obtained efficiently. The equation for the tangent can be found using the ideas of differentiation learned in the Mathematics II component of high-school mathematics. However, students in the humanities for whom differentiation and integration were not necessary for university entrance examinations may complain of not having learned these topics in school classes. In such a case a compromise may be reached with students by explaining a little advance knowledge regarding the 2nd order functions dealt with in Mathematics I.

2nd order functions are functions which are proportional to a 2nd order term. Students at junior high school learn about 1st order functions which are straight lines proportional in the 1st order, but in high school they learn about 2nd order curved-line functions which are proportional in the 2nd order. Having students draw diagrams of such 2nd order functions on note paper can be used to confirm that this has been learned.

I then put forward the problem of obtaining the equation for the tangent which touches such a 2nd order function

Problem 3. Find the equation of the tangent which touches $y = x^2$ at $x = 2$.

Many students complain that not having learned differentiation, they cannot use derivative based methods, so I explain a method of solving the problem using only Mathematics I knowledge without relying on differentiation. The 2nd order function is written $y = x^2$, and its tangent is written $y = ax + b$. The fact that they touch means that the 2nd order function and the straight line “intersect at a point” (Figure 3). Let us therefore seek the solution to the two equations according to the condition that it is a repeated solution (*i.e.*, there is only one solution).

$$x^2 = ax + b$$

$$x^2 - ax - b = 0$$

$$\text{From the discriminant, } D = a^2 + 4b = 0,$$

$$b = -\frac{a^2}{4}$$

The tangent passes through the point $P(x, x^2)$ on the 2nd order function, so substituting this into the equation for the tangent yields

$$x^2 = ax - \frac{a^2}{4}.$$

Solving this yields

$$4x^2 = 4ax - a^2,$$

$$a^2 - 4xa + 4x^2 = 0,$$

$$(a - 2x)^2 = 0$$

The constants in the equation for the tangent may thus be determined.

$$a = 2x$$

$$b = -\frac{4x^2}{4} = -x^2$$

Compiling the results above, the equation for the tangent which intersects the function at $P(x, x^2)$ may be expressed in terms of X and Y as follows.

$$Y = 2xX - x^2$$

Substituting $x = 2$ into this equation yields

$$Y = 4X - 4.$$

The equation of the tangent line is thus

$$y = 4x - 4.$$

This calculation can certainly be made without using differentiation.

5. The Newton-Raphson Method

The equation of the tangent was calculated above using Mathematics level I knowledge, so now let's try using differentiation to find the tangent. Firstly, let's try to find the gradient of the tangent to the 2nd order function $y = x^2$ at the point $P(x, y)$. The gradient of the tangent may be found as the differential coefficient. The idea behind the differential coefficient is as follows. To begin with, it does not involve thinking of $P(x, y)$ as having an intersection at a single point, but rather, as intersecting two points such that the second has its x coordinate offset by a small distance h , $Q(x+h, (x+h)^2)$. The slope between P and Q may then be calculated.

$$\frac{\Delta y}{\Delta x} = \frac{(x+h)^2 - x^2}{(x+h) - x} = \frac{2hx + h^2}{h} = 2x + h$$

Since the intersection does in fact occur at a single point, Q is brought close to P . That is to say, Δx is brought close to 0. When Δx tends to 0,

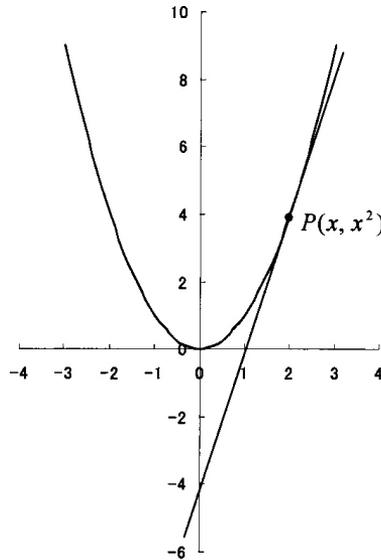


Figure 3: 2nd order function and its tangent

Δy also approaches 0. You might think that since both the numerator and the denominator are close to 0 the value of the fraction could not be obtained, but it may in fact be properly calculated. This kind of manipulation is known as a limit. This value is the gradient of the tangent, and the derivative function is expressed as $f'(x)$.

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{h \rightarrow 0} \frac{2hx + h^2}{h} = 2x$$

Also, since the tangent passes through the point $P(x, x^2)$, expressing the equation of the tangent in terms of the coordinates (X, Y) yields

$$Y - x^2 = 2x(X - x)$$

$$Y = 2xX - x^2,$$

This is the same as that obtained using the discriminant method.

Well then, the focus of this chapter was the calculation of root 2, and in the case of root 2, the 2nd order function

$$y = x^2 - 2$$

is used. This is just the original 2nd order function shifted along the Y -axis by -2 . The equation of the tangent to this 2nd order function can be obtained. The calculation is easy because the derivative function can be used. In general, the equation of the tangent is

$$Y - f(x) = f'(x)(X - x).$$

and substituting $f(x) = x^2 - 2$ and $f'(x) = 2x$ into this equation yields

$$Y - (x^2 - 2) = 2x(X - x).$$

Rearranging this,

$$Y = 2xX - x^2 - 2.$$

Now then, let's use this formula to find the value when the tangent cuts the X -axis, *i.e.*, when $Y = 0$.

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

The value $x = x_1$ may be used temporarily in the right-hand side of this equation. Substituting this into the right-hand side and representing the value calculated as $X = x_2$ yields a recurring equation with the following form.

$$x_2 = \frac{1}{2}\left(x_1 + \frac{2}{x_1}\right).$$

The initial value of x_1 may be set for example as $x_1 = 1$. This value is substituted into the equation above, and the value of x_2 is calculated. Then for the next candidate for the value of x_1 , the calculated value of x_2 replaces the original value of x_1 . If this process is repeated a number of times, it is possible to calculate the value of root 2 by hand. Table 1 shows that the result obtained after 3 iterations of this calculation is 1.414215, which is correct to 5 decimal places. This method is known as the Newton-Raphson method, taking the name of the man who thought up the derivative method, Newton.

When roots are calculated in BASIC or C programs, it is sufficient to set the error tolerance at which to break off iterating the calculations so as to stopping when the absolute value of the difference between x_1 and x_2 falls below a certain value ϵ (*e.g.*, 10^{-7}).

| Number of iterations | x_1 | x_2 |
|----------------------|---------|----------------------|
| 1 | 1 | $3/2 = 1.5$ |
| 2 | $3/2$ | $17/12 = 1.416$ |
| 3 | $17/12$ | $577/408 = 1.414215$ |

Table 1. Calculation of root 2 using a recurring formula.

Problem 4. Obtain the value of root 5 to within 3 decimal places using the methods introduced for root 2.

Computers have built-in functions like SQRT, SQR, *etc.*, for calculating roots. Calculators have a button with the root symbol ($\sqrt{\quad}$) marked on it. Both computers and calculators perform iterative calculations according to the Newton-Raphson method, and do not contain root data prepared in advance and stored as a database.

The capabilities of personal computers and calculators have really come along. Between the 1960s and early 1970s we didn't have personal computers, calculators were very expensive and the technology was in the process of development. There was a difference in the time taken by the four arithmetic operations and the calculation of roots which could be confirmed with the naked eye. The result of an addition or subtraction appeared in an instant when the Enter key was pressed, but a root calculation took a few seconds. There were roots for which the result was revealed immediately, and roots which took time to calculate. Those were the good old days when you could stare at a busy calculator and think about the effort it was putting into the iterative calculations we have been discussing in this chapter.