University of Toronto
Mathematics Network

Question Corner and Discussion Area

This document reflects the contents of the Question Corner and Discussion Area *portion of the University of Toronto Mathematics Network's web site*

`http://www.math.toronto.edu/mathnet/questionCorner/`

as of April 19, 1999. Additional material may have been added to the site since then.

This page has been a place for high-school students and others to ask questions about mathematical topics and receive accurate and informative answers.

Unfortunately, due to limited resources we have been unable to make staff available to continue providing this service, though we hope to be able to provide it again in the future.

However, please browse the large selection of questions and answers already posted; you will likely find your question addressed there somewhere.

Question Topics

Questions Arranged Roughly By Subject

Infinite Sequences, Series, and Recursions

- An Infinitely Recurring Square Root
- Finding the Sum of a Power Series
- The Sum of the Geometric Series $1 + 1/2 + 1/4 + \cdots$
- A Complicated-Looking Infinite Sum
- A Geometric Series Arising From an Arithmetic Series
- Finding the Ratio from the Sum of a Geometric Series
- Numbers Defined By Infinite Sums
- A Sequence Describing a Bouncing Ball

Geometry–General

- Angle Between Vertices of a Tetrahedron
- Counting Obtuse Triangles in an Inscribed Polygon
- Fractals and their History
- Four-dimensional Pyramids

Geometry–Euclidean and Non-Euclidean

- Multiplying Matrices
- How To Graph The Inverse Of A Function
- Does Every Function Have an Antiderivative?
- Symmetry of Functions and their Derivatives
- Factorials of Non-Integral Values

Polynomial and Transcendental Equations

- Solution to the Transcendental Equation $2^x + 3^x = 5$
- Limit of the Sequence $a(n) = \cos(a(n-1))$
- Solving a Quadratic with Non-Constant Coefficients
- Solution to a Functional Equation

Number Systems and Number Theory

- Largest Possible Number?
- Calculating Digits of Pi in Other Bases
- Tetrahedral and 4-Tetrahedral Numbers
- Existence of Shapes with Irrational Dimensions
- What are the Origins of Number Systems?
- The Hypercomplex Numbers
- Why is the Product of Negative Numbers Positive?
- The Number Zero
- Why You Can't Divide Nine By Zero
- Why is $x^0 = 1$?
- How To Express A Repeating Decimal Number As A Fraction
- Published Proof of Fermat's Last Theorem
- The $n = 4$ Case of Fermat's Last Theorem
- The Case n=3 Of Fermat's Last Theorem
- A Question from the IMO
- How To Find The Least Common Multiple
- A Geometric Proof That The Square Root Of Two Is Irrational

Historical Questions

- Which U.S. President Re-Proved the Pythagorean Theorem?

Discussion Topics

Below is a list of discussion topics that were created when we had staff available to process discussion topics and question submissions.

- Is Deductive Geometry Worth Salvaging in the High-School Curriculum?
- Anyone Have Interesting Graphing Calculator Problems?
- How Much Does School Size affect Performance in Math Contests?
- Starting a Math Club
- Mathematical Communication
- Teaching Linear Equations With CDROM Technology
- Teaching Addition Using Fractions and Decimals Together
- Success in Mathematics and Future Success
- Student Misconceptions about Complex Numbers
- Motivating Students

Contents

An Infinitely Recurring Square Root

Asked by Gopikrishna Srinivasan on Friday Nov 24, 1995:

> Hi!
>
> My name is Krishna. Before when I was in the math club, they posed a question. The question is:
>
> what is the 'sqrt[1+sqrt 1+sqrt 1 .... ?
>
> The square root is abbreviated with 'sqrt'. Also note that the sqrt sign is within the sqrt sign and so on.
>
> Thank you.
>
> Krishna.

One good way to tackle this problem is to ask the following question:

> Assuming this number exists, is there an equation that it must satisfy, an equation which is simple enough that it can be easily solved?

If we let x denote this number, notice that we have

$$x = \sqrt{1 + \underbrace{\sqrt{1 + \sqrt{1 + \cdots}}}_{x}}$$

and the part underlined with a brace is the same thing as x itself. This means that x must satisfy the equation $x = \sqrt{1+x}$ which can be solved by squaring both sides to get $x^2 = 1 + x$, using the quadratic formula to find

$$x = \frac{1 \pm \sqrt{5}}{2}.$$

Since $(1 - \sqrt{5})/2$ is negative, but x is positive, x has to be the other root, namely $(1 + \sqrt{5})/2$.

What this method does is it tells you that, *if* such a number x exists, then you can figure out what it has to be.

That's probably all you were asking for, but strictly speaking it isn't a complete answer. It leaves open the question: Does this number exist at all?

To prove that it does, you need some ideas from calculus: *every bounded, increasing sequence has a limit.* Here we are looking for the limit of the sequence

$$\sqrt{1}, \sqrt{1 + \sqrt{1}}, \sqrt{1 + \sqrt{1 + \sqrt{1}}}, \ldots.$$

Can you show that this sequence is bounded and increasing? (Hint: you will need to use mathematical induction. Boundedness is the trickiest one to prove; try proving by induction that all of the terms are less than 2. Use the fact that the nth term $a(n)$ and $(n-1)$st term $a(n-1)$ are related by

$$a(n) = \sqrt{1 + a(n-1)}$$

to show that, if $a(n-1)$ is less then 2, then $a(n)$ must be also).

Post another question if you want more of an answer on this part!

Finding the Sum of a Power Series

Asked by Khanh Son Lam, student, College de Maisonneuve on January 24, 1998:

> Hi!
>
> My question is about geometric series. I read about the one that you solved, but this one is a little bit different :
>
> What is the sum from $i = 0$ to infinity of $(x^i)(i^2)$?
>
> Thanks.

The series you have described is not a geometric series. It is an example of a more general class of series called *power series*, which are of the form $\sum_{n=0}^{\infty} a_n x^n$ where the coefficients a_n don't depend on the variable x. In your example, $a_n = n^2$.

A key fact about power series is that, if the series converges on an interval of the form $|x| < R$, then it "converges uniformly" on any closed subinterval of that interval. I won't attempt to explain what that means, but will mention instead an important consequence (which is not always true for series that are not power series): *the series can be integrated and differentiated term by term*, in the sense that, if you define

$$f(x) = \sum_{n=0}^{\infty} a_n x^n,$$

then

$$f'(x) = \sum_{n=0}^{\infty} n a_n x^{n-1}.$$

This means that, if you start with the geometric series $\sum_{n=0}^{\infty} x^n$ which is known to converge to $1/(1-x)$ when $|x| < 1$ (as described in the answer to another question), the following is true for all $|x| < 1$ by differentiating both sides of the equation:

$$\frac{1}{(1-x)^2} = \sum_{n=0}^{\infty} n x^{n-1}.$$

If you multiply both sides by x you get something close to what you want:

$$\frac{x}{(1-x)^2} = \sum_{n=0}^{\infty} n x^n.$$

Differentiating both sides again and multiplying by x again gives you what you want:

$$x\frac{1+x}{(1-x)^3} = \sum_{n=0}^{\infty} n^2 x^n.$$

Therefore, your series converges to $(x+x^2)/(1-x)^3$, provided $|x| < 1$. (If $|x| > 1$, it diverges).

This particular technique will, of course, work only for this specific example, but the general method for finding a closed-form formula for a power series is to look for a way to obtain it (by differentiation, integration, etc.) from another power series whose sum is already known (such as the geometric series, or a series you can recognize as the Taylor series of a known function).

Most series don't have a closed-form formula, but for those that do, the above general strategy usually helps one to find it.

The Sum of the Geometric Series $1 + 1/2 + 1/4 + \cdots$

Asked by Krishna Srinivasan on Friday Dec 22, 1995:

> My name is Krishna. I'm now in Grade 12. When I was in Grade 11, I saw a question in the math club.
>
> Actually, I have already asked a similar question like the one below, but this one is little different.
>
> The question is,
>
> What is the value of, 1 + (1/2) + (1/4) + (1/8) ...... ?
>
> How can you find a definite value for an answer when it keeps on continuing?
>
> Thank you.
>
> Krishna

The reason an infinite sum like $1 + 1/2 + 1/4 + \cdots$ can have a definite value is that one is really looking at the *sequence* of numbers

1
$1 + 1/2 = 3/2$
$1 + 1/2 + 1/4 = 7/4$
$1 + 1/2 + 1/4 + 1/8 = 15/8$

etc.,

and this sequence of numbers $(1, 3/2, 7/4, 15/8, \ldots)$ is converging to a limit. It is this limit which we call the "value" of the infinite sum.

How do we find this value?

If we assume it exists and just want to find what it is, let's call it S. Now

$$S = 1 + 1/2 + 1/4 + 1/8 + \cdots$$

so, if we multiply it by $1/2$, we get

$$(1/2)S = 1/2 + 1/4 + 1/8 + 1/16 + \cdots$$

Now, if we subtract the second equation from the first, the $1/2$, $1/4$, $1/8$, etc. all cancel, and we get $S - (1/2)S = 1$ which means $S/2 = 1$ and so $S = 2$.

This same technique can be used to find the sum of any "geometric series", that it, a series where each term is some number r times the previous term. If the first term is a, then the series is

$$S = a + ar + ar^2 + ar^3 + \cdots$$

so, multiplying both sides by r,

$$rS = ar + ar^2 + ar^3 + ar^4 + \cdots$$

and, subtracting the second equation from the first, you get $S - rS = a$ which you can solve to get $S = a/(1 - r)$. Your example was the case $a = 1$, $r = 1/2$.

In using this technique, we have assumed that the infinite sum exists, then found the value. But we can also use it to tell whether the sum exists or not: if you look at the finite sum

$$S = a + ar + ar^2 + ar^3 + \cdots + ar^n$$

then multiply by r to get

$$rS = ar + ar^2 + ar^3 + ar^4 + \cdots + ar^(n + 1)$$

and subtract the second from the first, the terms ar, ar^2, $\ldots$, ar^n all cancel and you are left with $S - rS = a - ar^(n + 1)$, so

$$S = \frac{a(1 - r^{n+1})}{1 - r}.$$

As long as $|r| < 1$, the term $r^(n + 1)$ will go to zero as n goes to infinity, so the finite sum S will approach $a/(1 - r)$ as n goes to infinity. Thus the value of the infinite sum is $a/(1 - r)$, and this also proves that the infinite sum exists, as long as $|r| < 1$.

In your example, the finite sums were

1 = 2 - 1/1
3/2 = 2 - 1/2
7/4 = 2 - 1/4
15/8 = 2 - 1/8

and so on; the nth finite sum is $2 - 1/2^n$. This converges to 2 as n goes to infinity, so 2 is the value of the infinite sum.

A Complicated-Looking Infinite Sum

Asked by Andrew Gill, student, State College Area High School on August 8, 1997:

> I have been trying to find the sum of:
>
> $$\sum_{q=0}^{\infty} \frac{(2q+1+\pi\sqrt{n})(\pi^2 n)^q}{(2q+1)!}$$
>
> (where n is a positive integer), but have not been able to find the process for evaluating this power series. Any and all information on this problem would be greatly appreciated.

This summation can be broken up and rewritten. Splitting the fraction into two pieces by writing $(2q+1+\pi\sqrt{n})(\pi^2 n)^q/(2q+1)!$ as the sum of $(2q+1)(\pi^2 n)^q/(2q+1)!$ and $(\pi\sqrt{n})(\pi^2 n)^q/(2q+1)!$, then rewriting $(\pi^2 n)^q$ as $(\pi\sqrt{n})^{2q}$ and simplifying, gives

$$\sum_{q=0}^{\infty} \frac{(\pi\sqrt{n})^{2q}}{(2q)!} + \frac{(\pi\sqrt{n})^{2q+1}}{(2q+1)!} = \sum_{k=0}^{\infty} \frac{(\pi\sqrt{n})^k}{k!}$$

(the first sum runs over all even integers $k = 2q$, the second over all odd integers $k = 2q+1$, so the combined sum runs over all integers k).

Now you should notice that the last summation is just the power series expansion of $e^{\pi\sqrt{n}}$.

A Geometric Series Arising From an Arithmetic Series

Asked by Buddy Jalmasco, teacher, Xavier School on June 27, 1997:

> How do we prove that $5^a, 5^b, 5^c, \ldots$ is a geometric sequence if it is known that $a, b, c, \ldots$ is an arithmetic sequence?

An arithmetic sequence is one for which the difference between consecutive terms is a constant, so $b - a, c - b, \ldots$ all have the same value (call it r).

A geometric sequence is one for which the ratio between consecutive terms is a constant. So, to show that $5^a, 5^b, 5^c, \ldots$ is a geometric sequence, just look at the ratios: $5^b/5^a = 5^{b-a}, 5^c/5^b = 5^{c-b}, \ldots$. These ratios are all equal to the same value, namely 5^r (because $b-a = c-b = \ldots = r$). Therefore, $5^a, 5^b, 5^c, \ldots$ is a geometric sequence (whose common ratio is 5^r where r is the common difference of the arithmetic sequence).

Here's another way to think of this. In general, the nth term of an arithmetic sequence is of the form $nr + s$ where r and s are two constants. So we can write $a = (0)(r) + s$, $b = (1)(r) + s$,

$c = (2)(r) + s$, and so on. The sequence $5^a, 5^b, 5^c, \ldots$ becomes $5(0r + s), 5(1r + s), 5(2r + s), \ldots$ where the nth term in the new sequence is given by $5(nr + s)$.

We can then rewrite the terms of the sequence to obtain $(5^s)((5^r)^0), (5^s)((5^r)^1), \ldots, (5^s)((5^r)^n), \ldots$ Note that the nth term of this sequence is of the form qp^n where $q = 5^s$ and $p = 5^r$ and therefore this is a geometric sequence (a geometric sequence is, by definition, one whose nth term takes the form qp^n for some constants p and q).

Finding the Ratio from the Sum of a Geometric Series

Asked by Ned Piburn on August 8, 1996:

> Thank you for maintaining such an interesting and useful web page. The following problem comes from a real life application: defining a smooth calculational grid for a computational fluid dynamics program.
>
> Consider the sum of N terms of a geometric series:
>
> $S = A1 + A1 * R + A1 * R^2 + ... + A1 * R(N - 1)$
>
> We know that this is equivalent to:
>
> $S = A1 * (1 - R^N)/(1 - R)$
>
> We can rearrange this equation to solve for N to get:
>
> $N = LN((S * R + A1 - S)/A1)/LN(R)$.
>
> The challenge is to solve for R. From the form of the first equation, we know that what we seek is one of the roots of a polynomial of degree $N - 1$. If we view the equation in physical terms we should satisfy ourselves that 1) there is a unique real solution and 2) there are reasonable restrictions that we can apply that will eliminate the unsought complex roots.
>
> The physical model that the algebra represents is: S is a distance to be spanned by N computational cells. The first cell has a width of $A1$ the second a with of $A1 * R$ and so on. What we seek is an expansion (or contraction) factor, R, that will cause the sum of N cell widths to be exactly S.
>
> From this we can state the following inequalities:
>
> | | |
> |---|---|
> | $S > 0$; | there must be a distance to span. |
> | $A1 > 0$; | the first cell must have width. |
> | $A1 < S$; | we won't span the distance with the first cell. |
> | $N > 1$; | we have to have at least a second cell to cover the remaining distance. |
> | $R > 0$; | follows from above. |
>
> From a purely intuitive view there ought to be a closed form generalized solution for R. But I have been unable to see my way to anything but an iterative numerical solution. Can you supply some insight?
>
> Thank you and best wishes, Ned Piburn

Closed-form solutions are more of a rarity than one might think, and not as important as one might think. The equation you are trying to solve can be written as

$$R^N + R(N - 1) + ... + R + 1 - q = 0 \text{ (*)}$$

where $q = S/A1$. This equation does not have a general closed-form solution for R in terms of familiar functions (like addition, multiplication, taking powers, and extracting roots) of N and q, unless $N < 5$.

(Incidentally, although the form $S = A1 * (1 - R^N)/(1 - R)$ looks simpler, it's not any easier to solve for R; because although, when you rearrange it into a polynomial equation $R^N - qR + q - 1 = 0$, it looks nicer than (*), it is simply $(R - 1)$ times (*). So in finding its roots the first thing you'd do is discover and discard the root $R = 1$; then you'd divide by $(R - 1)$ and be left with (*). So it's no easier to use this form; in fact, it's mathematically less convenient because you have the extra root $R = 1$ to worry about).

Solutions by radicals of polynomials of degree ≤ 4 were known by the 1500's (the degree 2 case being the familiar quadratic formula). For a couple of hundred years mathematicians tried and failed to find such a formula for quintic (degree 5) equations; eventually it was discovered that no such general formula exists, because associated to each such polynomial is an abstract mathematical object called a *Galois group*, and only when the Galois group has certain nice properties does a solution by radicals exist. All Galois groups for polynomials of degree ≤ 4 have this property, but the Galois groups for most degree ≥ 5 polynomials do not. Some do, but equation (*) is not among them (except for very special values of q like 0 or 1).

There are general solutions of quintic equations involving other functions like hypergeometric functions, but from a practical point of view that doesn't matter too much since the only way to evaluate them is by numerical calculations anyway. And besides, that's only for $N = 5$.

However, just because there isn't a "closed-form" solution doesn't mean R is not expressible as a function of N and q! In fact, your "physical arguments" can be translated into precise mathematics. The conditions you gave, namely that $N > 1$, $R > 0$, and $S > A1$ (so that $q > 1$), mean that R is uniquely defined as a function of N and q. The mathematical proof of that goes like this:

Let $f(R) = R^N + \cdots + R + (1 - q)$. This f is a continuous function, with $f(0) < 0$ and $f(q-1) > 0$. (Reason: $f(0) = 1 - q < 0$ since $q > 1$, and $f(q-1) = (q-1)^N + \cdots + (q-1)^2 + (q-1) + (1-q)$ which, since $(q-1)^N$, …, $(q-1)^2$ are all positive, is greater than $(q-1) + (1-q) = 0$.)

Therefore, by the intermediate value theorem, $f(R) = 0$ for some value of R in the range $0 < R < q - 1$.

Also, this R is unique: there can be no other $R > 0$ for whichs $f(R) = 0$, because f is an increasing function when $R > 0$ (its derivative is $NR^(N-1) + ... + 2R + 1$ which is positive).

Therefore, for each $N > 1$ and $q > 1$, there is a unique corresponding $R > 0$ for which (*) holds. The equation *does* define R as a function of N and q. You could even go on to describe properties of that function; you can prove that it's continuous and differentiable, for example.

It just happens that this function isn't representable as a combination of the few functions we're familiar with, so to calculate its values we need numerical approximations. But that's not an unexpected thing; very few functions are. Even the trigonometric and exponential functions cannot be written as combinations of the simpler functions; that's why we give new names to them. The same thing is true of the function representing R in terms of N and q; if you wanted to use it a lot, you'd just give that function a name and a notation, the way one did with functions like "sin" and "log". To calculate its values you use some kind of numerical method; that's exactly what one does when one calculates sines or logarithms.

Numbers Defined By Infinite Sums

Asked by a student at Mission Bay High School on September 21, 1997:

> What if someone did this : $k = 1/\sqrt{1} + 1/\sqrt{2}\ldots$ and it comes to some amount. Would that number be "special"? Are there any other special constants that haven't been discovered?

This particular sum does not add up to any (finite) number. (We say that the sum *diverges*; you can also think of it as "adding up to infinity").

To see this first note that $1/\sqrt{1} + 1/\sqrt{2} + 1/\sqrt{3} + \ldots$ is greater than $1/1 + 1/2 + 1/3 + \ldots$. The second sum is often called the *harmonic series* and is known to diverge. This is because we can regroup the terms and write it as $(1/1) + (1/2) + (1/3 + 1/4) + (1/5 + 1/6 + 1/7 + 1/8) + \ldots$. Now this sum is greater than $(1/1) + (1/2) + (1/4 + 1/4) + (1/8 + 1/8 + 1/8 + 1/8) + \ldots$. Note that each parenthesized group except the first adds up to $1/2$. So, after the first group, the sum is at 1. After the next two groups, the sum as at $1 + (1/2) + (1/2) = 2$. After the next two groups, the sum is at $2 + (1/2) + (1/2) = 3$. And so on. Continuing in this way, we see that the sum eventually passes every finite number, so the total sum is infinite.

In general there are many numbers which are defined by infinite sums. For instance $e = 1 + 1/2! + 1/3! + 1/4! + \ldots$ where $n! = (n)(n-1)(n-2)\cdots(2)(1)$. Pi can also be expressed as a sum, although it is a bit more complicated.

Your example can be modified to add up to a finite number. For each power $p > 1$, the sum $1/1^p + 1/2^p + 1/3^p + \ldots$ does add up to a finite value. But no special name is given to that value.

There are infinitely many real numbers (in fact, what mathematicians call "uncountably infinitely many" which is even more than just "infinitely many"!), so there are plenty of numbers that nobody has had occasion to think about in a special way. That doesn't mean they haven't been "discovered", though; after all, until now nobody has probably ever before written down the number

$$2374348627846827468726482678346278348236$$

but that doesn't mean I somehow "discovered" it by writing it down for the first time. A "discovery" would be discovering that some number had special, unique, and important properties that no other number has.

It is difficult to say which numbers are "special." Numbers such as π and e are special because they occur naturally in many situations in applied mathematics. There are a variety of other constants in mathematics which are somewhat less useful but nevertheless still relevant to some people.

A Sequence Describing a Bouncing Ball

Asked by a student at SRF K-OAC High School on December 11, 1997:

> In my math class, our teacher has given us an independent study on sequences and series. This is one question for example: A superball bounces to 3/4 of its initial height when dropped on dry pavement. If the ball is dropped from a height of 16 metres,
>
> a)How high does it bounce after the fifth bounce?
>
> b)How far does the ball travel by the time it hits the ground for the sixth time?
>
> Please help me answer this question as soon as possible. This may be an easy question for you since you probably are used to dealing with complex questions and intelligent students but could you please just help me out with this one? Thank you!
>
> Respectfully,

> Melanie Graine
>
> Grade 11 student at K-OAC High School

A good place to start is to think of how high the ball goes on each bounce. This gives you a sequence: $a(0)$ is height from which ball was dropped (16), $a(1)$ is the height of the first bounce, which is 3/4 of $a(0)$, $a(2)$ is the height of the second bounce, and so on. Your questions are to (a) find $a(5)$, and (b) find the sum of $a(0)$ (amout it falls initially) plus $2a(1)$ (amount it goes up and down after hitting the ground the first time), plus $2a(2)$, and so on, up to $2a(5)$.

The information you have about the sequence is that each bounce is 3/4 the height of the previous bounce. This tells you that $a(n+1) = 3/4a(n)$. So you can now calculate $a(1) = (3/4)a(0) = (3/4)16 = 12$, $a(2) = (3/4)a(1) = (3/4)(12) = 9$, and so on. Then you can answer the questions asked.

More interesting are the questions "how long will the ball keep bouncing, and how far will it go in total?" Even if it makes an infinite number of bounces, the total time taken, and the total distance travelled, will be finite! To answer these questions you need a formula for $a(n)$, which you can get by writing

$$a(n) = (3/4)a(n-1) = (3/4)^2 a(n-2) = ... = (3/4)^n a(0) = (3/4)^n 16$$

Then the total distance travelled can be found by taking the sum of this infinite series (actually, you need to take twice the sum less the first term, since on all bounces except the initial drop the ball has to go both up and down). The way you find the sum of an infinite series like this is described in the answer to another question; the answer turns out to be exactly 112 metres.

The fact that the total time taken is finite requires some knowledge of physics to calculate the time travelled for each bounce, but you end up with a similar series that has a finite sum.

Angle Between Vertices of a Tetrahedron

Asked by Lee Lude, student, Michigan City High School on February 5, 1998:

> Given a regular tetrahedral with a point in the center, find the angle formed from this center point to two corners (next to each other) in the tetrahedral. In chemistry terms you are actually proving that the the electrons in a tetrahedral shaped molecule are 109.5 degrees apart. No where in the proof however can you use the 109.5 degrees. This is the angle that you want to find.

This souds like an assignment question. Our goal here is to help you understand mathematical concepts (especially ones that go beyond the standard curriculum), so we would prefer you to ask "Here is a question I have been given, and I would like to understand better how I should go apart tackling such-and-such a part of it ..." rather than simply asking us the same question that was asked of you.

There are several ways to tackle this question. The simplest involves using vectors. Think about the four vectors from the centre of the tetrahedron to the four corners. You know several things about these vectors:

1. These vectors all have the same length l.
2. The angle θ between any pair of vectors is the same, and from the theory of dot products one has the fact that $\mathbf{u} \cdot \mathbf{v} = ||\mathbf{u}|| ||\mathbf{v}|| \cos\theta = l^2 \cos\theta$ (where $\mathbf{u}$ and $\mathbf{v}$ are any two of the vectors).

3. By symmetry, the vector sum of all four vectors is zero.

Combining these pieces of information, if $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$, and $\mathbf{x}$ are the four vectors, you have

$$\mathbf{0} = \mathbf{0} \cdot \mathbf{u} = (\mathbf{u} + \mathbf{v} + \mathbf{w} + \mathbf{x}) \cdot \mathbf{u} = (\mathbf{u} \cdot \mathbf{u}) + (\mathbf{v} \cdot \mathbf{u}) + (\mathbf{w} \cdot \mathbf{u}) + (\mathbf{x} \cdot \mathbf{u}) = l^2 + 3l^2 \cos\theta$$

and therefore $\cos\theta = -1/3$, which you can solve for θ to get an answer of approximately 109.5 degrees.

Counting Obtuse Triangles in an Inscribed Polygon

Asked by Krishna Srinivasan, student, Stephen Leacock C.I. on June 16, 1997:

> A regular 18-sided polygon is inscribed in a circle, and triangles are formed by joining any 3 of the 18 vertices. How many obtuse triangles are there?

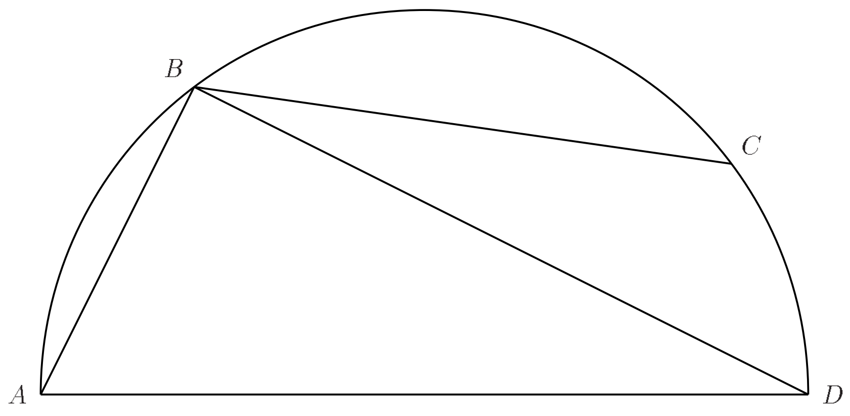
Is this a question *you* have, or is it a question from an assignment or project that you are asking for help on? If it is an assignment or project question, you should really be asking us "here is a problem I am trying to solve and can't. I've tried ..., but I'm stuck on ...; where can I go from here?" rather than simply asking us the same question that was asked of you. That helps us to answer more appropriately. It also makes it more likely that we will answer at all, for normally we don't answer questions that read as if they might be taken directly from an assignment or test, but we're always happy to answer questions *about* such problems and the concepts behind them. (In this case, though, the question is sufficiently non-standard that we will go ahead and answer it).

The real problem is that the question is ambiguous. There are two completely different possible interpretations, and therefore two completely different possible answers. We'll address each.

But first, here's an important characterization of obtuse angles: *Suppose point B is on the arc of a circle going from A to C. Then angle ABC is obtuse if and only if the arc is less than a semicircle.*

You can see this in two ways. First of all, there's a theorem in geometry which says that angle ABC is half the measure of the "other" arc from A to C (the one not containing B). Therefore, angle ABC is obtuse if and only if the other arc is more than 180 degrees, which is true if and only if the arc from A to C through B is less than 180 degrees (less than a semicircle).

You can also see it by remembering that any angle inscribed in a semicircle is a right angle. Now, if the arc from A to C through B is less than a semicircle, you can complete it into a semicircle ending at point D as shown below. Then angle ABD is a right angle, and it is clear from the picture that angle ABC is greater than angle ABD and therefore obtuse. This shows that if the arc is less than a semicircle the angle ia obtuse; you can draw similar pictures which show that, if the arc is more than a semicircle, the angle is acute.



Now here are the two interpretations of the question, and their solutions:

Interpretation #1. *You choose any 3 of the 18 vertices, call them A, B, and C, and join them. How many obtuse triangles are formed by this process?*

First, here's a convenient way to describe the relative positions of the vertices. Let's label the vertices A, B, and C in order clockwise around the polygon. Starting at A and moving clockwise, count how many edges there are from A to B, from B to C, and from C to A. The total of these three numbers will always be 18, since by the time you get back to A you have covered all 18 edges.

For example, if A is at vertex 1 of the 18-gon, B is at vertex 7, and C is at vertex 11, the three numbers will be 6, 4, and 8.

The triangle ABC is obtuse if and only if A, B, and C all lie on some arc which is less than a semicircle (i.e., encompasses fewer than 9 vertices of the polygon). This is the same as saying that two of the three numbers add up to less than 9, which is equivalent to the third number being greater than 9.

Therefore, triangle ABC is obtuse if and only if one of the three numbers is greater than 9.

(In fact, you can say more than this. Suppose the three numbers are (a, b, c). The angle at vertex B equals half the measure of the arc from C to A that doesn't include B. This arc encompasses c edges. Each edge corresponds to an angle of $360/18 = 20$ degrees, so the arc from C to A is $20c$ degrees, so angle B is half that, namely $10c$ degrees. Similarly, angle C is $10a$ degrees and angle A is $10b$ degrees. For example, if the three numbers are 3, 5, and 10, that means the angles of triangle ABC are 50, 100, and 30 degrees. This shows that ABC is obtuse if and only if one of the numbers is greater than 9).

Now the only question is, how many other triangles are there? You get an extra triangle each time one of the lines you drew combines with two polygon edges to form a triangle. This happens every time you draw a line between two vertices that are separated by only 2 edges: in other words, it happens every time one of the three numbers is a "2".

You should be able to easily see that each of these "extra" triangles is obtuse (one of its angles is the angle between two edges of the 18-gon, which is 160 degrees).

Therefore, the number of obtuse triangles equals the number of "extra" triangles (which is the number of 2's among the three numbers), plus 1 if the "main" triangle ABC is obtuse (which happens if one of the three numbers is a 9).

You should be able to convince yourself that this means the number of obtuse triangles equals the number of 2's and 9's among the 3 numbers, and that this can be anywhere from 0 to 3, depending on how the vertices A, B, and C were chosen.

Interpretation #2. *For any possible choice of three vertices A, B, and C, you draw the triangle ABC. Thus you have drawn 816 triangles in all (that's the number of ways of choosing 3 out of 18 things). How many of these 816 triangles are obtuse?*

Label the vertices of each obtuse triangle A, B, and C, in order clockwise, starting at the vertex *before* the obtuse angle (so that angle B is the one which is obtuse).

Let's ask ourselves, "Suppose we know where vertex A is. Where could vertices B and C be?" In order for angle B to be obtuse, the clockwise arc from A through B to C must be less than a semicircle, so vertex C must come before the 9th vertex after A.

Therefore, B and C could be anywhere among the next 8 vertices after A. The number of possible choices for B and C equals the number of ways of choosing a pair of vertices out of 8 possibilities.

This means that for each of the 18 possible positions for vertex A, there are x different obtuse triangles with A at that position, where x is the number of ways of choosing two out of eight things.

You should be able now to figure out the rest of the answer.

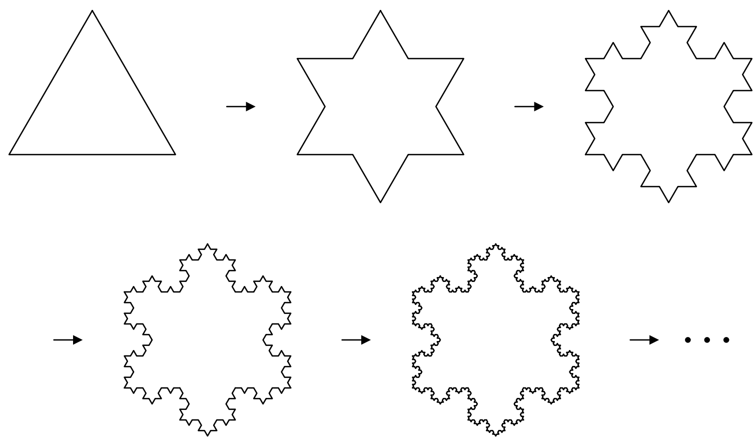
Fractals and their History

Asked by Anuradha (last name unknown) on July 2, 1997:

> Please assist me in learning about "fractals." What are fractals? How are they applied in mathematics at secondary school levels and later at higher levels of school or education?

The name "fractal" arises from the concept of a fractional dimension. What this means exactly is a difficult to say in simple language and we will simply try to give a feel for what fractals are and what sorts of behaviors they exhibit.

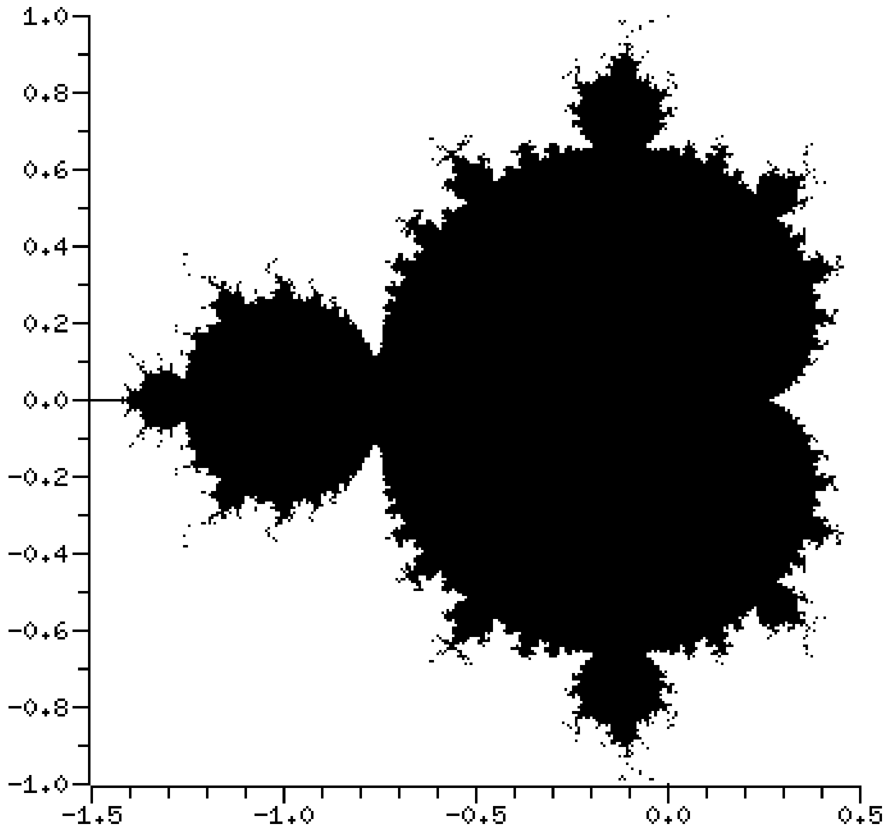
Generally fractals are the result of some which is procedure repeated again and again. One of the most basic examples of a fractal is obtained in the following way. Start off with an equilateral triangle which has sides of length 1. Now on each edge of the triangle, add a new equilateral triangle with sides of length 1/3. Now in the middle of each side of this new shape, add a triangle with sides of length 1/9. Continue this process, each time adding new triangles to each side which are 1/3 the size of the triangles added in the last stage. When you are done (infinitely many steps later!) you have the desired fractal.



Now let us take a moment to examine some of the properties of this strange new "shape." The area of this object can be calculated by adding up the area which we added at each stage. It is not difficult to see that this sum is actually a geometric series which converges to some finite area. One can also check that, after adding the new triangles at some stage, the perimeter of the shape is four-thirds what it used to be. Thus after repeating the process n times, the perimeter of our shape is $(3)(4/3)^n$. After repeating this process infinitely many times to get our fractal, the "perimeter" becomes infinite. Some of you may be uneasy about this last part since it is not really easy to tell how to define the perimeter of such an object. It turns out though that there is a way to define a notion of perimeter for such a shape and that we come to the same conclusion when things are handled with a bit more care.

Another interesting thing to note is that if we magnify our fractal along its edge, it looks the same, no matter how large the magnification is. This is another property typical of fractals.

A more well known fractal, the Mandelbrot set, is a little harder to describe mathematically. It's definition is based on the multiplication of the complex numbers. Start with a complex number z_0. From it define $z_1 = (z_0)^2 + z_0$. Assuming that we know what z_n is, define z_{n+1} to be $(z_n)^2 + z_n$. The points in the Mandelbrot set are all those points which stay relatively close to the point $0 + 0i$ (in the sense that they are always within some fixed distance of $0 + 0i$) as we repeat this process. As it turns out, if z_n is ever outside of the circle of radius 2 about the origin for some n, it won't be in the Mandelbrot set. When you see colored pictures of the Mandelbrot set, what you are really seeing are the points outside the set. The color code corresponds to how many iterations it takes to place these points outside this circle of radius 2.



It turns out that fractals occur frequently in nature. This is often due to how nature performs its own iterations. For instance, as a tree grows, new branches grow from old branches. As these branches mature, they sprout new branches of their own. It is possible to implement a relatively simple algorithm on a computer and have it produce a remarkably accurate rendition of a tree. The shore line of a lake or large body of water also tends to exhibit fractal behavior. One can tell by comparing aerial photos of shore lines to the contours of certain fractals whether or not the shore is natural.

Asked by Jill (last name unknown), student on December 16, 1996:

> What can you tell me about the history of fractals?

Fractals have their roots in 19th century mathematics. In Jean Perrin, *Les Atoms*, and William Fellers' book *Introduction To Probability*, they discussed both real and simulated Brownian motion (a natural phenomenon which is chaotic in nature).

In 1918 both Fatou and Julia worked on what we would think of as the more standard types of fractals.

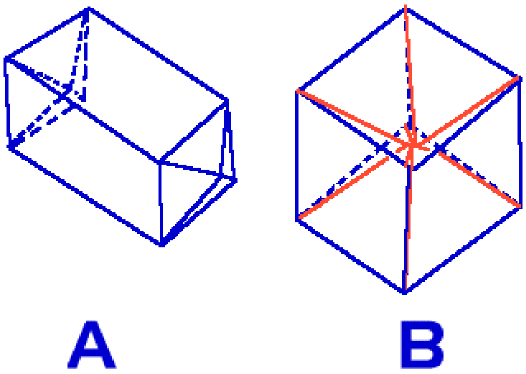
One of the most inspiring works however seems to have been Poincaré's *Vorlesungen uber die Theorie der Automorphen Funktoren* published in 1897 which contained many influential illustrations. His drawings of hyperbolic tesselations were embellished by M.C. Escher and made into a form of art which itself could be argued to be closely related to fractals.

Some of the modern interest in fractals among the general public comes from the computer-assisted work of the IBM fractal project and 20th century mathematicians like Benoit Mandelbrot after whom the Mandelbrot set is named. Pictures of this set are quite fascinating and show surprising features at every level of detail. Until the introduction of the computer, very little from this area of mathematics was in the form of graphics. IBM's fractal project added the pictures to what is inherently an incredibly visual field of mathematics. This added dimension of the field in turn gave rise to new discoveries and incarnations of fractals such as mountains and clouds.

Four-dimensional Pyramids

Asked by Anders Ericson, student, Sweden on November 12, 1997:

> Hi,
>
> My name is Anders Ericson and I am a Swedish Student on my 12th year (is that equal to college)
>
> I program a lot and have made a program that draws a 4D cube as well as a 5- and 6D cube. Now I'm wondering how a 4D pyramid looks like (an egyptian pyramid with a square bottom). I've come up with two alternatives that I have drawn in the attached GIF-file.
>
> Figure A is supposed to be a stretched cube with a pyramid cut out in the bottom and placed at the top. Figure B is a normal cube with lines from all the corners in to the middle, that means that its created by six normal pyramids.
>
> I think that it is the B figure that is right but it can also be some other figure.



Anders Ericson

Picture B would be a correct depiction of a 4-dimensional pyramid with a "square" bottom.

That is to say: if in 3 dimensions you have a pyramid consisting of a square bottom and an additional top vertex that is joined to all the vertices of this square, then the 4-dimensional analogue of this would be a figure that has a cubical bottom and an additional vertex that is joined to all the vertices of this cube. Your picture B depicts this.

The sides of this pyramid are 3-d pyramids, whose base is one of the square faces of the bottom cube. This is analgous to the fact that the sides of a 3-d pyramid are 2-d pyramids (triangles) whose base is one of the line segments of the bottom square.

I'm not entirely sure of the details of your picture A, such as what the dashed lines are meant to represent. However, it does not appear to be a correct description of the pyramid, because I cannot see any vertex with valence 8 (a vertex that has 8 edges emanating from it), which the true pyramid must have because the top vertex is adjacent to all 8 of the vertices of the bottom cube.

These square-based 3-d pyramids and cube-based 4-d pyramids are not actually the most natural mathematical objects. More natural is the 3-d tetrahedron (a pyramid whose base is a triangle rather than a square). The 4-dimensional analogoue of this is discussed in the answer to another question.

Euclidean Geometry

Asked by a student at Lincolin High School on September 24, 1997:

> What is Euclidean Geometry? Can you also give me an example of it. Thank you very much.

Euclidean geometry is just another name for the familiar geometry which is typically taught in grade school: the theory of points, lines, angles, etc. on a flat plane. It is given the name "Euclidean" because it was Euclid who first axiomatized it (rigorously described it).

Another reason it is given the special name "Euclidean geometry" is to distinguish it from non-Euclidean geometries (described in the answer to another question).

The difference is that Euclidean geometry satisfies the Parallel Postulate (sometimes known as the Fifth Postulate). This postulate states that for every line l and every point p which does not lie on l, there is a unique line l' which passes through p and does not intersect l (i.e., which is parallel to l).

Geometry on a curved surface, for example, may not satisfy this postulate, and hence is non-Euclidean geometry.

Euclidean Geometry in Higher Dimensions

Asked by Victor Humberstone on February 10, 1997:

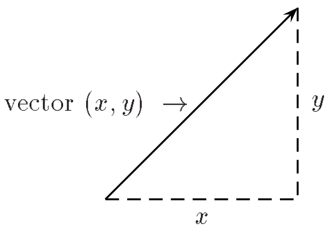
> I would like to know where I can find out a little more than high school maths on Euclidean Geometry. In particular, I would like to understand n-dimensional symmetrical 'solids' (esp 4, 5 dimensions.) My son has recently been asking about a drawing of a 'hypercube' (a 4-D cube) in an old book by George Gamov in which such an object was drawn and wants to understand how to extend the concept. I can't help! Can you help me to help him?

Euclidean Geometry in higher dimensions is best understood in terms of coordinates and vectors. In fact, it is these which even give *meaning* to geometric concepts in higher dimensions. So, let me start with a quick overview of those (which you, as a physics graduate, will know anyway and may want to skip, but others reading the page may not):

In 3 dimensions, we all have an intuitive understanding of what *length* and *angle* mean, and it is not at all clear how to extend these concepts to higher dimensions.

However, if we introduce coordinates and think about vectors, one can express both length and angle in terms of vector operations, and these operations readily generalize to any number of dimensions.

First of all, length is given by Pythagoras's Theorem: for a two-dimensional vector (x, y), its length is $\sqrt{x^2 + y^2}$, as illustrated by the right-angled triangle in the picture below.



The formula for the length of a 3-dimensional vector (x, y, z) is also easily obtained: you draw a right-angled triangle whose hypotenuse is the vector (x, y, z) and whose other two sides are $(x, y, 0)$ (whose length is $\sqrt{x^2 + y^2}$ since it's really just a 2-d vector) and $(0, 0, z)$ (whose length is just $|z|$). Applying Pythagoras's Theorem to this triangle gives

$$\text{length of } (x, y, z) = \sqrt{x^2 + y^2 + z^2}.$$

This formula for length readily generalizes to any number of dimensions. For example, a vector in 4-dimensional space can be given by four coordinates as (x, y, z, w), and its length is defined to be $\sqrt{x^2 + y^2 + z^2 + w^2}$ by analogy to the length formula in two and three dimensions.

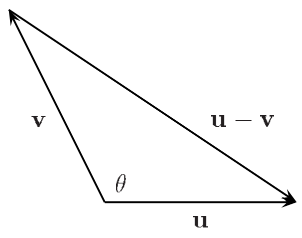
This, then, gives a definition of what the concept of *length* means in four and higher dimensions.

The other main geometric concept is that of *angle*. Again, we have an intuitive understanding of what it means in our 3-dimensional world, and it's not at first clear how to generalize it to higher dimensions.

However, if you apply the law of cosines to the three vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{u}-\mathbf{v}$ illustrated below, the angle θ between $\mathbf{u}$ and $\mathbf{v}$ satisfies the equation

$$|\mathbf{u}-\mathbf{v}|^2 = |\mathbf{u}|^2 + |\mathbf{v}|^2 - 2|\mathbf{u}||\mathbf{v}|\cos\theta$$

where $|\ |$ denotes the length of a vector.



Therefore, angle can be expressed in terms of length using

$$\cos\theta = \frac{|\mathbf{u}|^2 + |\mathbf{v}|^2 - |\mathbf{u}-\mathbf{v}|^2}{|\mathbf{u}|\,|\mathbf{v}|}.$$

Since the length formulas make sense in any number of dimensions, so too does the notion of angle, if defined using this formula.

Finally, both length and angle are expressible in terms of a fundamental vector operation, the *dot product*: the dot product of two vectors $(x, y, z, \ldots)$ and $(a, b, c, \ldots)$ is defined to be the number $xa + yb + zc + \cdots$ obtained by multiplying together corresponding components and adding. This definition works for any number of dimensions. For example, the dot product of the two 5-d vectors $(1, 2, 3, 4, 5)$ and $(6, 7, 8, 9, 10)$ is

$$(1, 2, 3, 4, 5) \cdot (6, 7, 8, 9, 10) = (1)(6) + (2)(7) + (3)(8) + (4)(9) + (5)(10) = 130.$$

Length relates to the dot product by the formula $|\mathbf{v}| = \sqrt{\mathbf{v}\cdot\mathbf{v}}$, and angle by the formula

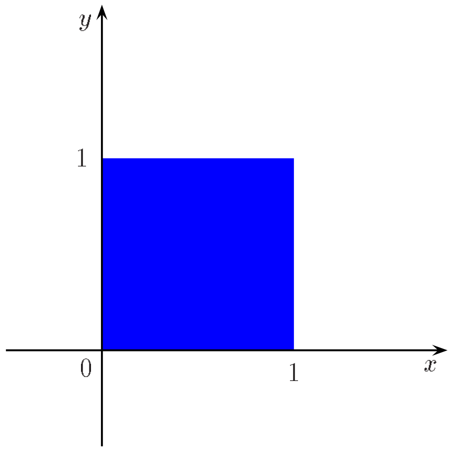
$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{|\mathbf{u}|\,|\mathbf{v}|} = \frac{\mathbf{u}\cdot\mathbf{v}}{\sqrt{\mathbf{u}\cdot\mathbf{u}}\sqrt{\mathbf{v}\cdot\mathbf{v}}}.$$

The dot product is the fundamental operation from which all the geometric concepts can be defined. In fact, this is how mathematicians study abstract and complicated geometric spaces: any time, in any context, that you have an operation which satisfies the basic algebraic properties that the dot product satisfies, that operation defines a notion of geometry on the things you are studying, and you can employ geometric principles to study it, even though it may at first glance have seemed a very un-geometric situation!

This is exactly how geometry in four-, five-, and higher-dimensional Euclidean spaces is defined.

Hypercubes

Shapes in higher dimensions can be expressed by giving equations among the coordinates, and a lot can be understood about them on that basis alone. For example, a 2-d square (with side length 1, located as shown below with respect to the coordinate axes) is the set of points (x, y) for which both x and y lie between 0 and 1.



A 3-d cube can be thought of as the set of points (x, y, z) for which x, y, and z each lie between 0 and 1.

In exactly the same way, a 4-d hypercube is the set of points (x, y, z, w) for which x, y, z, and w each lie between 0 and 1. The five-dimensional hypercube would be the set of points (x, y, z, w, u) for which x, y, z, w, and u each lie between 0 and 1. And so on: in n-dimensional space, the analogous thing to the hypercube is the set of points $(x_1, x_2, \ldots, x_n)$ such that each x_i satisfies $0 \leq x_i \leq 1$.

You can tell a lot about these objects just from that description. For example, let's ask ourselves, "what does the boundary look like?" We know that for an ordinary square, the boundary is just four line segments. For a cube, the boundary consists of six square faces. What about a hypercube?

Well, in order for a point (x, y, z, w) to be on the boundary, it means it's possible to move just a little bit away from that point and then be outside the hypercube: in other words, one of the coordinates would be less than 0 or greater than 1. The boundary case will be when one of the coordinates is exactly equal to 0 or 1.

There are eight possible such boundary conditions: for each of the four coordinates, there are two values (0 and 1) of that coordinate which give rise to a boudary piece.

This means that the boundary consists of eight pieces: the piece where $x = 0$, the piece where $x = 1$, the piece where $y = 0$, and so on.

What do each of these boundary pieces look like? Well, consider for example the piece $w = 0$. The remaining coordinates (x, y, z) are the only ones that are allowed to vary on the boundary, and they can range anywhere from 0 to 1. This is *precisely* the description of an ordinary three-dimensional cube.

Therefore, a hypercube's boundary consists of eight pieces, each of which is an ordinary cube.

If you carry out the same analysis in general, for n-dimensional space, you see that the n-dimensional hypercube has a boundary consisting of $2n$ pieces, each of which is an $(n-1)$-dimensional hypercube. (This is in complete agreement with what we already know about cubes and squares: the boundary of a cube is $(2)(3) = 6$ squares, the boundary of a square is $(2)(2) = 4$ line segments, and the boundary of a line segment is $(2)(1) = 2$ endpoints).

You can also ask questions like, "how many vertices does a hypercube have?" A vertex occurs whenever *all* the coordinates are either 0 or 1. If there are n coordinates, with 2 choices for each, that gives 2^n different vertices.

For example, when $n = 1$ we have a line segment, which has $2^1 = 2$ vertices. When $n = 2$ we have a square, which has $2^2 = 4$ vertices. When $n = 3$ we have a cube, which has $2^3 = 8$ vertices. When $n = 4$ we have a hypercube, which has $2^4 = 16$ vertices. And so on.

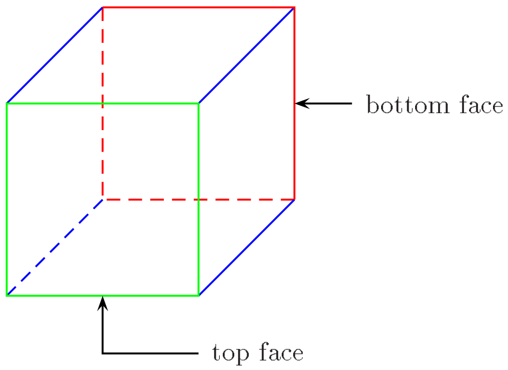
Visualization

Of course, even though you can tell a lot about a hypercube just by looking at formulas involving coordinates, it's still nice to be able to try to visualize it in some way. People have developed quite a lot of skill at drawing 2-dimensional pictures of 3-dimensional objects. In the same way, one can try to draw a 3-dimensional picture of a 4-dimensional object like the hypercube. Mathematically, this is called a *projection*; a function which takes each point in the 4-dimensional object and maps it onto 3-dimensional space. The problem is, when putting something on paper (or on a computer screen) you have to make a further projection onto 2-dimensional space! That makes it very hard to draw pictures which accurately convey the 4-d reality. I will try to do the best I can, though.

One way to project from four dimensions to three dimensions is to just ignore one coordinate. That is, the four-dimensional point (x, y, z, w) could be drawn simply as the three-dimensional point (x, y, z). Under this projection, a hypercube just looks like a cube.

To understand why this is so, think about looking at an ordinary cube with one eye (no depth perception) from directly above one of the faces. It would look just like a flat square. So, when projecting from 3-d to 2-d by ignoring one coordinate, a cube looks like a square. Similarly, when projecting from 4-d to 3-d by ignoring one coordinate, a hypercube looks like a cube.

A slightly better way to view a cube is to look at it, still from above, but somewhat diagonally; then you get a 2-dimensional picture like the one shown below:

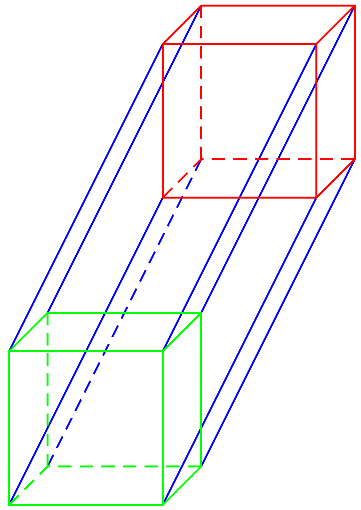


What's happened is the top and bottom faces of the cube are projected onto slightly different locations in the plane: the 2-d picture of the top face (an ordinary square, shown in green) is shifted down and to the left from the 2-d picture of the bottom face (another ordinary square, shown in red). Now you can also see the edges joining the top and bottom faces (shown in blue).

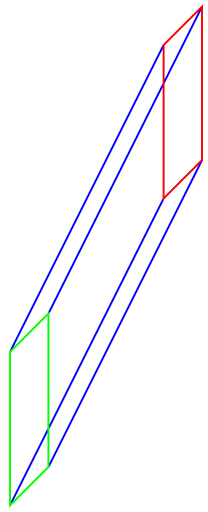
Mathematically: this corresponds to taking the 3-d point (x, y, z) and drawing, in 2-d, not the point (x, y), but rather, the point (x, y) shifted down and to the left by an amount that depends on z. That is, the point (x, y, z) is drawn as the 2-d point $(x - az, y - bz)$ (the scale factors a and b determine how much shifting to do). The above picture corresponds to $a = b = 1/2$; thus the points $(x, y, 0)$ on the bottom face are drawn at (x, y), resulting in a (red) square with x, y ranging from 0 to 1, while the points $(x, y, 1)$ on the top face are drawn at $(x - (1/2), y - (1/2))$, resulting in a (green) square in which x and y range from $-1/2$ to $1/2$.

You can do a similar thing with the hypercube: project it into 3-d, but shifting by an amount proportional to the fourth coordinate. That is, the 4-d point (x, y, z, w) is drawn at the 3-d point $(x - aw, y - bw, z - cw)$ for some constants a, b, and c.

When you do that, you get a cube corresponding to the bottom "hyperface", drawn in red below (*but remember what's drawn below is only a two-dimensional projection of the three-dimensional picture I'm talking about*), and a shifted cube corresponding to the top hyperface, drawn in green below. Now you can visualize the remaining edges of the hypercube, drawn in blue, by joining up each vertex of the top-hyperface cube with the corresponding vertex of the bottom-hyperface cube.

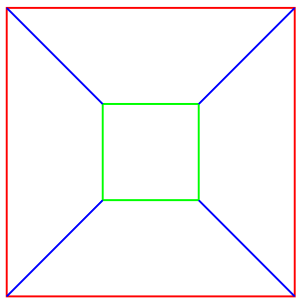


Remember I stated above that the hypercube's boundary consists of 8 pieces, each of which is a cube. (These are what I'm calling the "hyperfaces"). Well, two of them (the red and green cubes) are visible in the picture above. But what of the other six? It's hard to see them from this picture, but if you try hard, I hope you can see that the four rightmost blue edges, together with the four rightmost red edges and the four rightmost green edges, form a third cube.



The topmost blue, red, and green edges form the fourth boundary cube, the leftmost ones form the fifth, and the bottom ones form the sixth. The remaining two are harder to see, but the "frontmost" red square, the frontmost green square, and the blue edges joining them form the seventh boudary cube, thile the rearmost red square, the rearmost green square, and the blue edges joining them form the eighth one. To be able to see this well, you really have to do it from a 3-dimensional drawing; too much gets lost in the transition from 4-d to 3-d to 2-d to be able to see them clearly in the picture above.

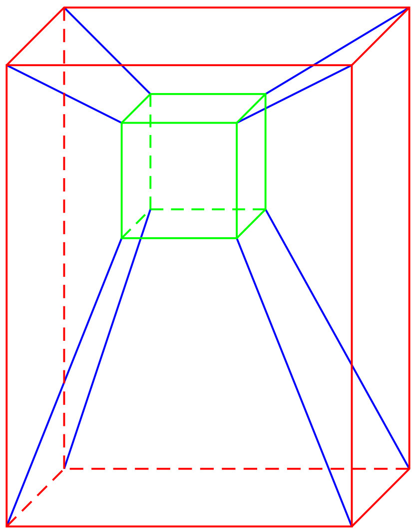
Another way to project a cube onto a plane is to shrink the top of the cube a little bit, so when you look at it from above you see something like



(to reconstruct the 3-d cube from this picture, think of pulling the inner square up from the paper, or out from the computer screen).

In other words, the top face of the cube is drawn in 2-d as a smaller square, inside the square with which the bottom face is drawn.

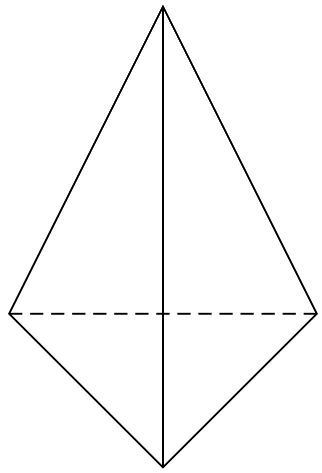
The analogous thing for hypercubes is to represent the top hyperface by one cube in 3-d and the bottom hyperface by a smaller cube inside the first cube. Then you can connect the corresponding vertices of the two cubes just as in the previous kind of projection. This is shown below.



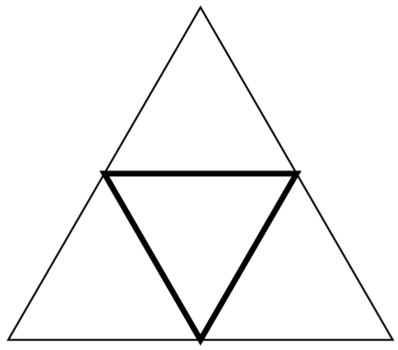
The visualization of higher-dimensional hypercubes is much harder. You could do a similar kind of thing (e.g., project a 5-d hypercube into 4-d as one 4-d hypercube inside another with corresponding vertices joined, then draw each of *those* hypercubes as a cube inside another), but the pictures get very complicated very quickly and I'm not even going to attempt a drawing!

Another 4-d solid: the 4-dimensional tetrahedron

Another symmetric 4-d "hypersolid" is the 4-dimensional analogue of a regular tetrahedron. Recall that a tetrahedron is a pyramid whose base and each of its sides are equilateral triangles:



One way to visualize such a tetrahedron in 2-d is to think of cutting it open along the three edges leading up to the peak, then laying it flat. This gives the following picture:



Observe that, in the picture, you have one triangle in the centre, then three other triangles, one attached to each of its sides. The pyramid is obtained by folding up those three triangles into the third dimension.

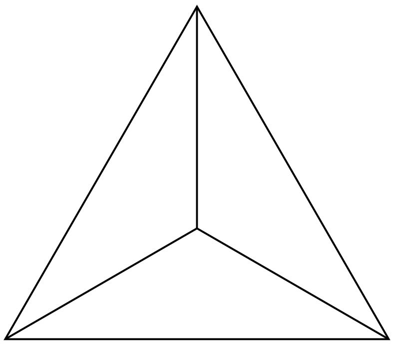
The 4-dimensional version can be visualized in the same way. You start with a solid tetrahedron in 3-d. To each of its four faces, you attach another tetrahedron. Now you've got a solid which it's a little hard to draw in 2-d, so I won't attempt it, but you might want to try making it with models:

1. First cut out of a piece of cardboard five copies of the picture above.
2. Fold along the three thick lines, until the tips of three outer triangles meet at the top, forming a pyramid. Attach them together with tape or glue.
3. Now you have five tetrahedrons. Choose one of them. Onto each of its four faces, glue one of the other four tetrahedrons.

Now you have one central tetrahedron with four others sticking out of it. The *4-dimensional tetrahedron* is the hypersolid obtained by bending each of those sticking-out tetrahedra (in four dimensions, you have an extra dimension to bend them into) until they meet at a point. Obviously, you cannot do this in three dimensions, but you can in four.

The 4-dimensional tetrahedron can also be visualized in a manner similar to the second method described for visualizing the hypercube.

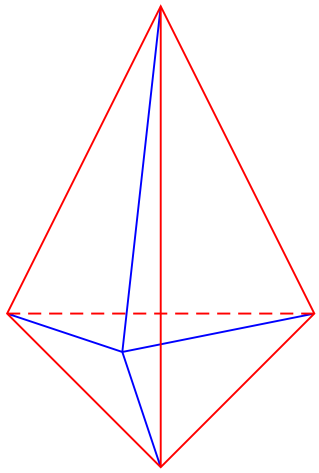
First lets understand how to visualize an ordinary 3-d tetrahedron in 2-d. If you stare directly at the top of it, with just one eye and no depth perception so that everything looks 2-dimensional, you see a picture like



where the peak of the tetrahedron is drawn inside the triangle representing the bottom face. This is a pretty good 2-d picture of a tetrahedron; you just have to imagine reaching in and pulling that central dot up and out of the page to get the real thing.

Observe that this picture consists of (a) a triangle, (b) a dot in the centre of the triangle, and (c) edges running between that dot and each of the triangle's three vertices.

In the same way, a 4-d tetrahedron can be visualized using the following 3-d model, illustrated below: (a) a tetrahedron (coloured red in the picture), (b) a dot in the centre of the tetrahedron, and (c) edges running between that dot and each of the tetrahedron's four vertices (coloured green).



There are analogues of the tetrahedron in every dimension, constructed by similar methods.

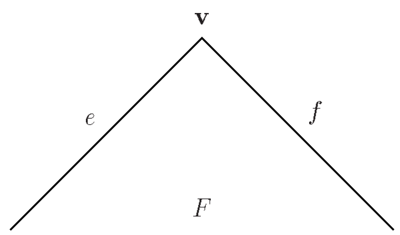
Other hypersolids

You ask a very interesting question when you ask about other symmetrical solids in 4 and 5 dimensions. The well-known mathematician Donald Coxeter, part of the mathematics department at the University of Toronto, spent many years of his life precisely on this question, worked out the answer, and wrote extensively about it. I will not here go into the details of how one finds all the different higher-dimensional regular solids, but I hope that the above examples (of higher-dimensional analogues of the cube and the tetrahedron, which exist in every dimension) are of interest to you and to your son.

The reason the question is so interesting is that the theory of regular shapes changes as you change dimension. In 2 dimensions, you can have a regular shape (a polygon whose lengths and angles are all equal) with *any number of sides*. So, there are an infinite number of regular polygons: the equilateral triangle, the square, the regular pentagon, the regular hexagon, and so on.

However, in three dimensions, there are only *five* different kinds of regular solids, instead of an infinite number! They are the tetrahedron (a 4-sided solid whose faces are all equilateral triangles), the cube (a 6-sided solid whose faces are all squares), the octahedron (an 8-sided solid whose faces are all equilateral triangles), the dodecahedron (a 12-sided solid whose faces are all regular pentagons), and the icosahedron (a 20-sided solid whose faces are all equilateral triangles). There aren't any others.

The reason is quite interesting. Look at one of the vertices of a regular polyhedron (call it v). A certain number of faces (let's say, k faces) come together at v. The number k has to be at least 3 (because, as illustrated below, each face F has two edges, e and f, touching v, and F must be adjacent to some other face along e and also to some other face along f, making for a total of at least three faces touching v).



However, there is also an upper limit on how large k can be, because the angles at any vertex must add up to less than 360 degrees (the reason is given below).

There are k of these angles. So, if the faces are equilateral triangles, in which the angles are 60 degrees, that means k times 60 must be less than 360, so $k < 6$. Therefore, the only possible regular polyhedra with equilateral triangles for faces are (a) one in which 3 faces meet at every vertex (this is the tetrahedron), (b) one in which 4 faces meet at every vertex (this is the octahedron), and (c) one in which 5 faces meet at every vertex (this is the icosahedron).

If the faces are squares, in which the angles are 90 degrees, we must have $k \times 90 < 360$, so $k < 4$. Thus $k = 3$ is the only possibility, and the only possible regular polyhedron with squares for faces is one in which 3 faces meet at every vertex; this is the cube.

If the faces are pentagons, in which the angles are 108 degrees, we must have $k \times 108 < 360$, so $k < 10/3$, so $k = 3$ is the only possibility. Therefore, the only possible regular polyhedron with pentagons for faces is one in which 3 faces meet at every vertex; this is the dodecahedron.

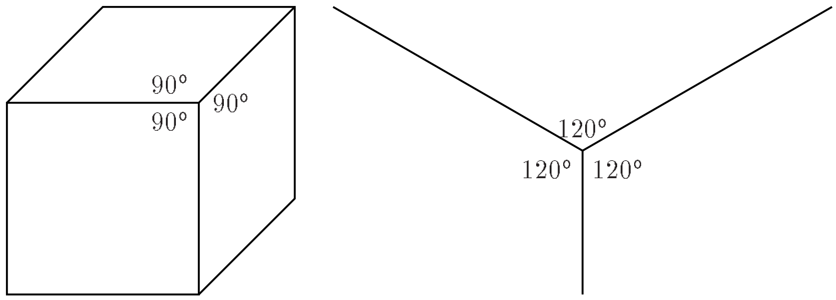
If the faces are polygons with 6 or more sides, the angles are 120 degrees or more, and we cannot have $k \times 120 < 360$ unless $k < 3$, which can't happen because there have to be at least three faces meeting at every vertex. Therefore, there is no regular polyhedron whose faces are hexagons, or any other polygon with 6 or more sides.

Now, what about 4 and higher dimensions? You can do a similar kind of analysis using solid angles instead of angles, and you find that there is only a limited, finite number of possibilities for the number of hyperfaces meeting at each vertex, depending on what the hyperfaces are (and, since the hyperfaces have to be regular solids, there are only five possibilities for those).

However, even if a particular combination is allowable according to the above analysis, there's no guarantee that there might not be some other reason ruling out the existence of such a regular hypersolid. So the work involved in figuring out exactly what kinds of regular polyhedra exist in 4 and higher dimensions is much more complicated than the analysis given above.

Nevertheless, I hope the above discussion gives you some idea of the kind of issues involved.

Finally, let me indicate why it is that the sum of the angles at a vertex of a regular polyhedron must add up to less than 360 degrees. If you take a two-dimensional picture of what happens at a vertex, as if you were looking through the vertex (directly toward the centre of the solid) without any depth perception, then in the two-dimensional picture the angles are larger than they are in the 3-d reality of the solid, and they add up to exactly 360 degrees. For instance, if you look at a corner of a cube, the real angles on the cube are 90 degrees, but the flattened out angles are 120 degrees:



This means the original angles in the solid must add up to less than 360 degrees.

The proof that the flattened-out angles are always less than the original ones uses vector algebra. If $\mathbf{u}$ and $\mathbf{v}$ are the flattened-out vectors corresponding to two of the edges, then the original 3-d vectors in the solid are these vectors plus a displacement vector $\mathbf{w}$ directed toward the centre of the solid.

Let a be the cosine of the flattened-out angle, and b the cosine of the original angle in the solid. If we scale things so that $\mathbf{u}$ and $\mathbf{v}$ have length 1, a is just $\mathbf{u}\cdot\mathbf{v}$. However, b is

$$\begin{aligned}
\frac{(\mathbf{u}+\mathbf{w})\cdot(\mathbf{v}+\mathbf{w})}{|\mathbf{u}+\mathbf{w}|\,|\mathbf{v}+\mathbf{w}|} &= \frac{\mathbf{u}\cdot\mathbf{v}+\mathbf{u}\cdot\mathbf{w}+\mathbf{v}\cdot\mathbf{w}+\mathbf{w}\cdot\mathbf{w}}{\sqrt{(\mathbf{u}\cdot\mathbf{u}+2\mathbf{u}\cdot\mathbf{w}+\mathbf{w}\cdot\mathbf{w})(\mathbf{v}\cdot\mathbf{v}+2\mathbf{v}\cdot\mathbf{w}+\mathbf{w}\cdot\mathbf{w})}} \\
&= \frac{a+|\mathbf{w}|^2}{1+|\mathbf{w}|^2} \quad (\text{since } \mathbf{u}\cdot\mathbf{v}=a \text{ and } \mathbf{u}\cdot\mathbf{w}=\mathbf{v}\cdot\mathbf{w}=0) \\
&= a+(1-a)\frac{|\mathbf{w}|^2}{1+|\mathbf{w}|^2}
\end{aligned}$$

which is greater than a (since $a<1$). So, the original angle in the cube has a larger cosine than the flattened-out angle, and hence it's a smaller angle.

Non-Euclidean Geometry

Asked by Brent Potteiger on April 5, 1997:

> I have recently been studying Euclid (the "father" of geometry), and was amazed to find out about the existence of a non-Euclidean geometry. Being as curious as I am, I would like to know about non-Euclidean geometry. Thanks!!!

All of Euclidean geometry can be deduced from just a few properties (called "axioms") of points and lines. With one exception (which I will describe below), these properties are all very basic and self-evident things like "for every pair of distinct points, there is exactly one line containing both of them".

This approach doesn't require you to get into a philosophical definition of what a "point" or a "line" actually is. You could attach those labels to any concepts you like, and as long as those concepts satisfy the axioms, then all of the theorems of geometry are guaranteed to be true (because the theorems are deducible purely from the axioms without requiring any further knowledge of what "point" or "line" means).

Although most of the axioms are extremely basic and self-evident, one is less so. It says (roughly) that if you draw two lines each at ninety degrees to a third line, then those two lines are parallel and never intersect. This statement, called *Euclid's Parallel Postulate*, seems more like a theorem than an obvious and self-evident property, and for centuries people tried and failed to prove it from the other axioms.

Eventually it was discovered that it is independent of the other axioms, in the sense that it is logically self-consistent to have some things called "lines" and other things called "points" which satisfy the other axioms but don't satisfy the parallel postulate. Any such a collection of things is called a non-Euclidean geometry.

There are many examples. Most concretely, if you do geometry on a curved surface instead of on a flat plane (where now "line" refers to the shortest path between two points, which obviously will not be straight if you are on a curved surface), you typically end up with a non-Euclidean geometry.

We already have on this web site a detailed description of one kind of non-Euclidean geometry called *projective geometry*. You can refer to that description for more details. It also includes a more precise description of Euclid's parallel postulate and the other axioms of geometry.

Followup question by Brent Potteiger on April 10, 1997:

> As a followup to my earlier question on non-Euclidean geometries, I would like to know how many types of non-Euclidean geometries there are. If possible, I'd also like to know a bit about each, and some source where I can find information about non-Euclidean geometries. Thanks a bundle!!!!!!!

In one sense there are infinitely many types. If you take a surface, then any way in which you choose to bend it will give you a different geometry. You could take any surface, for example, a sphere, or an ellipsoid (football-shaped surface), or a hyperboloid (hourglass-shaped surface), or any variation of any of these by bending in any way you like, and you will get a different geometry (the "lines" will be the paths of shortest distance between two points, even though they're not straight lines in the usual sense of the word). Hyperbolic geometry is probably the most important of these.

You can also extend geometric concepts to any number of dimensions.

Mathematically, any function which assigns a non-negative number to each pair of points determines a geometry: you just consider the distance between two points P and Q to be the whatever the function value $f(P,Q)$ is. There are a few restrictions on the function, like the fact that the distance between a point and itself has to be zero, the distance between two distinct points has to be greater than zero, the distance from P to R has to be less than the sum of the distance from P to Q and the distance from Q to R, and the distance between two points must not change too much if you move one of the points slightly. However, there are still infinitely many functions that satisfy these properties, so there are infinitely many such geometries. Geometries defined in this way are called Riemannian Geometries.

I don't know offhand of any reference that would be suitable for your level and that you'd be able to find in your school library. If I find one I will let you know.

Followup question by Noah Potvin, student, East Meadow high School on November 26, 1997:

> Hi, I'm back. I know you mentioned axioms on your website but what exactly are they? What is the fifth postulate and also please explain absolute geometry?I would also appreciate if you would explain hyperbolic geometry too. I know it's a bundle of questions but thanks for answering them.

An axiom is something which is held to be true as the basis or starting point for a logical argument, rather than something that is proven true by the argument. You can't prove things from nothing; you have to start with certain basic assumptions then derive the desired consequences from those assumptions.

Euclid's axioms were of two types: five he called "axioms", being basic principles of logic and reasoning (such as, if two things are each equal to a third thing then they are equal to each other), and five he called "postulates", being basic principles of geometry. His five postulates were (paraphrased in modern language):

1. for every pair of points, it is possible to construct a line segment joining them;
2. every line segment can be extended indefinitely in a straight line in either direction;
3. for every pair of points, it is possible to construct a circle centred at one point and passing through the other;
4. any two line segments emanating from the same point determine an angle. There's a definition of what it means for this angle to be a "right angle", and any two right angles are equal to each other.
5. the fifth postulate is much less self-evident. For a description, see our web page http://www.math.toronto.edu/mathnet/questionCorner/projective.html

It is impossible to prove things about points and lines without first knowing what points and lines are—or at least, knowing something about the nature of points and lines, even if the ambiguous philosophical question of what a point or a line actually *is* remains unanswered. Therefore, certain basic characteristics need to be assumed without proof. These characteristics are given by Euclid's postulates. (Modern mathematicians would tend to call them "axioms" rather than "postulates", for as language has evolved over time the word "axiom" in mathematics has come to mean what "postulate" used to mean).

If it seems unsatisfying to think of having to assume certain things without proof before you can prove other things, you can think of it in the following alternative way: the postulates give a definition of what one *means* by the words "point" and "line". These words mean any things that behave in the manner described by the postulates. Theorems in geometry then take the form

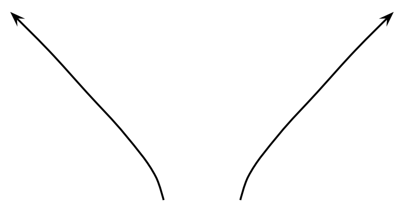
> if you have any collection of things called "points" and any other collection of things called "lines", and if they have the properties given in the postulates, then they will also have the property that ...

(followed by whatever the theorem in question is). In other words, the theorems in geometry are all logical consequences of the postulates, and don't depend on anything else about the nature of what points and lines are.

I am unfamiliar with what you mean by "absolute geometry."

Hyperbolic geometry is geometry in which, instead of Euclid's fifth postulate (which, roughly speaking, says that two lines starting out in the same direction will remain parallel to each other, staying just as far away from each other as they were at the start), has the property that lines starting out in the same direction will get further and further away from each other. The opposite extreme, where lines starting out in the same direction get closer to each other and eventually meet, is called elliptic geometry.

Hyperbolic geometry is similar in spirit (though different in detail) to geometry on a hyperboloid, an hourglass-shaped surface given by an equation like $ax^2 + by^2 - cz^2 = 1$ (where $a, b, c > 0$). If you start with two side-by-side vertical lines on the neck of this hourglass, and continue them as straight as you can while staying on the surface, the lines will end up getting farther and farther apart, as indicated below. That's the same kind of thing that happens in hyperbolic geometry.



Followup question by Kyle Shau, student, Conard High School on January 6, 1998:

> In Euclidean Geometry, isn't a line means a straight line and a plane means a flat plane? What are the definitions of "LINE," "PLANE," and "SPACE" in both Euclidean Geometry and Non-Euclidean Geometry?

The question is, what do you mean by "straight line"? To put it more fundamentally, what do you mean by "point"? The early Greeks realized that they didn't have any satisfying answer to that question; typical attempts went something like "a point is something that is infinitely small", but how do you know such things actually exist when you can't actually draw them?

That is why the axiomatic approach was adopted: rather than defining points and lines by some philosophical definition of what a "straight line" actually is, they are defined by what their properties are.

Under any axiomatic approach, be it Euclidean or non-Euclidean, a "geometry" is defined to be any set of things together with any collection of subsets of this set, that satisfies various properties. The "points" of the geometry are the elements of the set, and the "lines" of the geometry are the subsets.

Those are the definitions of "points" and "lines" in any form of axiomatic geometry.

Your intuitive notion of "straight line" and "flat plane" cannot be precisely defined (after all, if you draw a so-called "straight line" on a piece of paper, it really has some thickness and some ink smudges, not to mention the spaces between the atoms and molecules that make up the ink and the paper). So instead, one either opts for the axiomatic approach, or else moves to *analytic* *geometry*, whereby a point on the plane is defined to mean an ordered pair of numbers, and a line is defined to be a set of numbers (x, y) satisfying and equation of the form $ax + by = c$ where a, b, and c are constants.

Understanding Projective Geometry

Asked by Alex Park, Grade 12, Northern Collegiate on September 10, 1996:

> Okay, I'm just wondering about the applicability of projective and affine geometries to solving problems dealing with collinearity and concurrence. I'd really like to learn more on the topic, but I'm having trouble finding a book that gives the axioms of them both in a way that I can understand it. As far as I can understand it, there are no such things as parallel lines in projective geometry. How does that work? An explanation would be appreciated.

There are several different ways to think about geometry in general and projective geometry in particular.

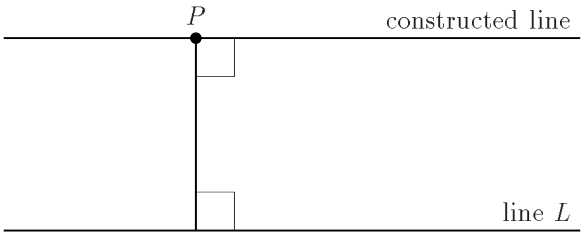
1. The axiomatic approach

This approach requires no philosophical definition of what a point or a line actually "is", just a list of properties (axioms) that they satisfy. The theorems of geometry are all statements that can be deduced from these properties. In this approach, the theorems of geometry are guaranteed to be true no matter what concept of "point" or "line" is being used and no matter how they are defined, as long as they satisfy the basic axioms.

Euclid wrote down a list of these axioms: five of them (though actually there are some other axioms implicit in Euclid's definitions). He called them *postulates*. The first four postulates are so self-evident that they clearly ought to be satisfied by anything worthy of the name "geometry". There are different ways of stating them and I don't know which you have seen, so I won't list them, but they basically say things like "for any pair of points, there is a unique line passing through both of them".

However, the fifth postulate isn't quite in the same category. Euclid's version of it was quite complicated; a simpler, equivalent version says that for any line L and a point P not on L, there exists a unique line that is parallel to L (never meets L) and passes through P. For this reason, the fifth postulate is called the *parallel postulate*.

At first glance it would seem that the parallel postulate ought to be a theorem deducible from the other more basic postulates, rather than something that has to be assumed separately. For centuries mathematicians tried to prove it, but always failed. For example: it follows from Euclid's first four postulates that there is a perpendicular line segment from P to L. Then you can draw a unique line through P that is perpendicular to that line segment. In some sense this is parallel to L because the two angles in the picture below are right angles, but how can one *prove* from this that this means the lines never intersect? We know they don't in our familiar mental picture of what an infinite flat plane looks like, but is that fact a *logical necessity* deducible from the other postulates?



Eventually it was discovered that the parallel postulate is logically independent of the other postulates, and you get a perfectly consistent system even if you assume that parallel postulate is *false*. This means that it is possible to assign meanings to the terms "point" and "line" in such a way that they satisfy the first four postulates but not the parallel postulate. These are called *non-Euclidean geometries*. Projective geometry is not really a typical non-Euclidean geometry, but it can still be treated as such.

In this axiomatic approach, *projective geometry* means any collection of things called "points" and things called "lines" that obey the same first four basic properties that points and lines in a familiar flat plane do, but which, instead of the parallel postulate, satisfy the following opposite property instead:

> *The projective axiom*: Any two lines intersect (in exactly one point).

(Depending on how one words the other axioms, they may need some slight modification too).

Using only this statement, together with the other basic axioms of geometry, one can prove theorems about projective geometry. Many of them are the same as ordinary geometry; the big difference is that there is no such thing as a pair of parallel, non-intersecting lines in projective geometry.

One interesting fact is worth mentioning: in projective geometry, points and lines are completely interchangeable! That is, any statement about points and lines would still be true even if you replaced all occurrences of the word "point" with the word "line", and vice versa. For instance, the basic axiom that "for any two points, there is a unique line that intersects both those points", when turned around, becomes "for any two lines, there is a unique point that intersects (i.e., lies on) both those lines", which is the property described above. There is a complete duality between points and lines in projective geometry.

Now, if this approach were all there was to projective geometry, it would be little more than an intellectual curiosity. All it means is that it logically consistent for there to be concepts called "points" and "lines" that satisfy the axioms of geometry with the projective axiom in place of the parallel postulate. It says nothing about whether such concepts would be interesting, relevant, or have any relation whatsoever to the normal concepts of lines and planes in Euclidean geometry.

However, there are other approaches that reveal the connection:

2. Euclidean Geometry Plus A "Line At Infinity"

Another way to approach projective geometry is to define it as follows:

Take each line of ordinary Euclidean geometry and add to it one extra object called a "point at infinity". Do this in such a way that the same extra object is added to parallel lines (so that the extended lines now intersect), while different extra objects are added to non-parallel lines (so that the extended lines don't intersect more than once).

In other words:

- To each line l of Euclidean geometry, associate some other object $f(l)$, in such a way that $f(l) = f(l')$ if and only if l and l' are parallel.

  [There are lots of ways to do this. For example, you could let $f(l)$ be the slope of l (a real number, or the symbol "∞" if l is vertical). Alternatively, you could let $f(l)$ be the counterclockwise angle from some fixed reference line to l. The precise method you use is unimportant.]

- The *points* of projective space are the points in Euclidean geometry together with these additional objects $f(l)$ (which are called the *points at infinity*).

- The *lines* of projective space are lines l in Euclidean space together with the extra object $f(l)$ attached. In addition, the collection of all the extra objects together is also called a *line* in projective space (called the *line at infinity*).

This definition satisfies all the axioms of projective geometry. For example, here's a proof that any two of these "lines" L and L' intersect in exactly one "point":

> If one of L and L' (say, L) is the line at infinity and the other (L') is not, then they intersect at exactly one point because by definition L' contains exactly one point at infinity.
>
> consists of an ordinary Euclidean line together with one point at infinity. That is, we can write $L = l \cup \{f(l)\}$ and $L' = l' \cup \{f(l')\}$ where l and l' are Euclidean lines.
>
> If l and l' intersect at a point p then $f(l)$ does not equal $f(l')$ (since $f(l)$ only equals $f(l')$ when l and l' are parallel), so p is the one and only intersection point of L and L'.
>
> If l and l' do not intersect, then $f(l) = f(l')$ (since l and l' are parallel), so again L and L' have exactly one intersection point.

This view of projective geometry makes it relatively easy to answer questions of concurrence and collinearity.

For example, what does a collection of concurrent lines in projective space look like? It is one of three things:

1. a collection of the projective extensions of lines that were concurrent in Euclidean space, or

2. a collection of the projective extensions of lines that were parallel in Euclidean space (but are now concurrent at a point at infinity), or

3. case 2 above together with the line at infinity.

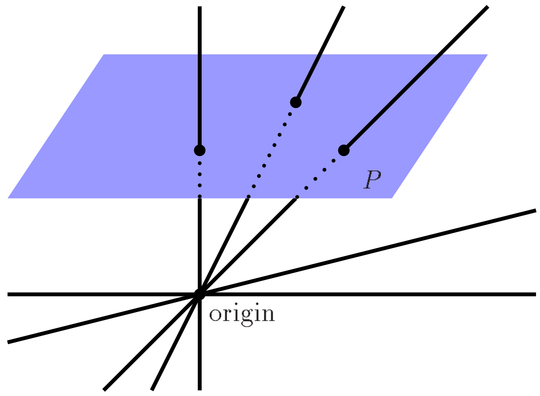
As for collinearity: points that are collinear in Euclidean space are still collinear in projective space. Also, the points at infinity are all collinear in projective space. A point p at infinity is only collinear with a collection of finite points if $p = f(l)$ where l is the Euclidean line through the finite points.

Although this view of projective geometry helps answer your question, it's still a little artificial, with all this talk of just "adding extra objects at infinity". There are two other, much more natural, ways of looking at it.

3. Lines In Space

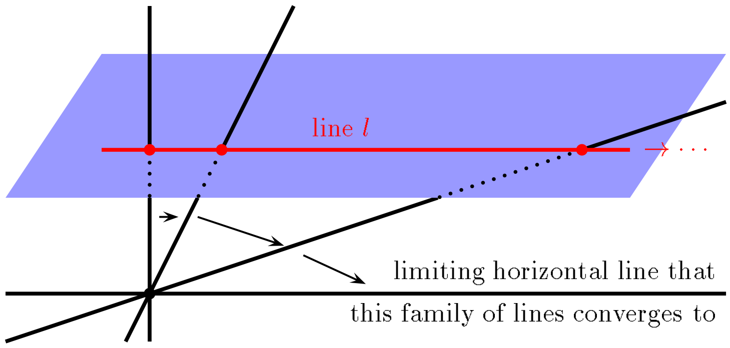
Projective geometry can be thought of as the collection of all lines through the origin in three-dimensional space. That is, each *point* of projective geometry is actually a *line through the origin* in three-dimensional space. The distance between two points can be thought of as the angle between the corresponding lines. A *line* in projective geometry is really a *family of lines through the origin* in three-dimensional space.

To see how this ties in with the previous view of projective geometry, let P be a horizontal plane in space that does not pass through the origin. As can be seen in the picture below, every line through the origin passes through exactly one point on P, except for the horizontal lines.



So there is a one-to-one correspondence between the points on the ordinary plane P, and some of the points in projective space (namely, all non-horizontal lines through the origin in 3-d space). The remaining points in projective space are horizontal lines through the origin in 3-d space; these are the “points at infinity”.

Now think about a line l in P. This corresponds to a family of lines through the origin in 3-d space, as shown. As you move to infinity on the line l, the corresponding lines through the origin actually converge to a horizontal line parallel to l, so this limiting line should be included in the family.



You can see this explicitly if you look at the direction vectors of the lines. Suppose that the plane P is at height 1 above the origin in 3-d space, and you have a line l given by $y = mx + b$ on P. Then a typical point p on l has 3-d coordinates $(x, mx+b, 1)$. This means that $(x, mx+b, 1)$ is a direction vector for the line through the origin that passes through p. The unit direction vector is obtained by dividing this vector by its length, to get

$$\frac{(x, mx+b, 1)}{\sqrt{1+x^2+(mx+b)^2}}.$$

It is an exercise in limits to show that this converges to $(1, m, 0)$ as $x \to \infty$; in other words, this family of lines converges to a horizontal line through the origin with slope m. Notice that the intercept b does not appear in the limit: parallel lines l, l' in P (ones with the same m and different b) correspond to families of lines through the origin in 3-d space which converge to the *same* horizontal line through the origin.

What all this means is that, in projective space, the "line" corresponding to l is actually a family of lines through the origin consisting of: (1) the lines that pass through l, and (2) the limiting horizontal line. Using the language in which a line through the origin in 3-d space is called a "point" in projective space, and a horizontal one is called a "point at infinity" in projective space: *the line in projective space corresponding to the Euclidean line l on P consists of (1) the points in projective space that correspond to the points on l, and (2) the point at infinity in projective space that is parallel to l.*

This is the same as what we saw before: lines in projective space consist of lines in Euclidean space with an added point at infinity. The difference is that now there's a definite geometric interpretation of the points at infinity. They're not just artificially added fabrications; rather, they're horizontal lines through the origin in 3-space. The remaining, non-horizontal lines through the origin in 3-space are in one-to-one correspondence with the points in a standard Euclidan plane P.

4. A Non-Orientable Surface

There's yet another way to understand projective geometry: it is the geometry of curves on a rather weird surface. This surface cannot be embedded in the standard 3-dimensional world we live in; you'd need to live in 4 dimensions to visualize it completely. Nevertheless, it has the following fairly easy interpretation: take a sphere (spheres in mathematics are hollow surfaces, not solid balls), and think of gluing together all pairs of antipodal points so that they become the same point. (Antipodal points are those which are on the most opposite part of the sphere from each other. For instance, if the centre of the sphere is at the origin, then the antipodal point of (x, y, z) is $(-x, -y, -z)$.)

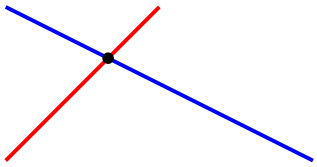
Another way to think of it is to just take the top hemisphere, and then "seal it up" into a closed surface by gluing each point on the equator to its opposite point.

Now, if you put a sphere with its centre at the origin in 3-d space, then every line through the origin passes through exactly 2 antipodal points in the sphere (and therefore exactly one point in our surface, after those two antipodal points get glued onto the same spot). So, there is a one-to-one correspondence between points in projective space (lines through the origin in 3-d space) and points on this weird surface.

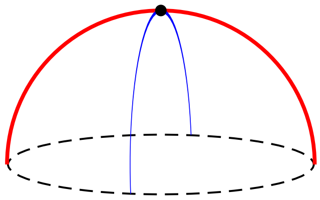
What is a "straight line" on a surface like a sphere or our glued surface? It's not really straight, obviously, but it can still be defined as giving the shortest distance between two points. On a sphere, the shortest path between two points is an arc that is part of a "great circle" on the sphere (one whose centre is the centre of the sphere). So, the "straight lines" on a sphere are the great circles.

The points on the top hemisphere of the sphere (excluding the equator) correspond to the points in a standard Euclidean plane as shown in the picture below. The points on the equator are the "points at infinity". Intersecting lines in the standard Euclidean plane are great circles that intersect as shown on the sphere:

In Euclidean Space:

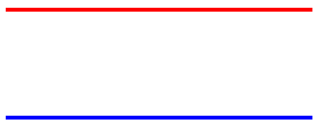


On the Sphere:

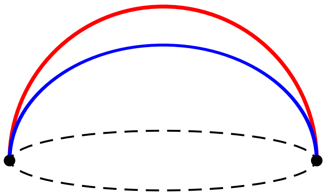


Parallel lines in the standard Euclidean plane are great circles that intersect only at the equator, as shown below:

In Euclidean Space:



On the Sphere:



(note that although these great circles intersect in two points, they become the same point after antipodal points get glued together so it's still true that in projective space, any two lines intersect in exactly one point).

In summary, then, projective geometry can be thought of as the study of points and "lines" (great circles) on a surface obtained by gluing the equator of a hemisphere to itself. It's very similar to the study of points and "lines" (great circles) on a sphere, except that on a sphere every pair of lines intersects in exactly two (antipodal) points instead of one.

Vectors in Projective Geometry

Asked by Mark Pimm-Smith on August 28, 1997:

> How do you define and operate with vectors in projective geometry?

(Note: the interested reader will find a brief introduction to projective geometry in response to another question on this web site).

Vectors in projective geometry—and in any kind of geometry on a curved space, for that matter—arise as *tangent vectors*. The best example to think of is the surface of a sphere. At any point on the sphere there is a tangent plane, and this tangent plane is a "vector space": a set of vectors that can be added and scaled as desired. Vector techniques can thus be used on the sphere. For example, given two intersecting curves on the sphere, one can measure the angle between them by looking at their tangent vectors and finding the angle between these vectors (with the dot product).

The same thing is true for the "projective plane" in projective geometry, and also for any of a wide class of mathematical objects called *manifolds*. Roughly speaking, a manifold is any set of points which can be covered by "patches", each of which can be flattened into something that looks like part of ordinary Euclidean space of some dimension. A sphere is an example of a manifold; the top and bottom hemispheres can be flattened into disks, even though the entire sphere itself cannot be. The projective plane is also an example of a manifold, as are most other smooth shapes you can think of.

On each manifold, there is a tangent space associated to every point. When the manifold is embedded in ordinary Euclidean space (as a sphere is), the tangent space is easy to visualize (in the case of a sphere, it's just a plane sitting inside ordinary space). But even in more abstract cases when things aren't as easy to visualize, the tangent space still exists, and vector techniques can be used to say a lot about the geometry of the manifold (such as defining the angle between two curves, even defining the length of a curve as the limiting value of successive approximations by short tangent vectors). All of this comes about through the techniques of calculus, which can relate tangent vectors of curves (which are essentially a kind of derivative) back to the curve itself.

I realize the above description is very sketchy (it's also a little inaccurate, for a knowledgeable reader will see that I have left out some important conditions and technical details). This is a huge subject, known as *differential geometry*, and it is hard to give a comprehensive introduction to it in such a short space. The interested reader may want to consult the book *Calculus on Manifolds* by Michael Spivak, which gives a reasonably elementary introduction to the subject (but it still requires some multivariable and advanced calculus as a prerequisite).

Do Parallel Lines Meet At Infinity?

Asked by a student at St-Joseph Secondary School on October 5, 1997:

> Could you help me prove that parallel lines meet at infinity or that infinity begins where parallel lines meet.
>
> I am curious. Could this ever happen?

The answer to the question depends on exactly what kind of geometry you are dealing with and what "points" and "lines" mean.

If you are talking about ordinary lines and ordinary geometry, then parallel lines do not meet. For example, the line $x = 1$ and the line $x = 2$ do not meet at any point, since the x coordinate of a point cannot be both 1 and 2 at the same time.

In this context, there is no such thing as "infinity" and parallel lines do not meet.

However, you can construct other forms of geometry, so-called non-Euclidean geometries. For example, you can take the usual points of the plane and attach to them an additional point called "infinity" and consider all lines to also include this additional point. In this context, there is a single "infinity" location where all lines meet. In a geometry like this, all lines intersect at infinity, in addition to any finite point where they might happen to meet.

Or, you could attach not just one additional point, but a whole collection of additional points, one for each direction. Then you can consider two parallel lines to meet at the extra point corresponding to their common direction, whereas two non-parellel lines do not intersect at infinity but intersect only at the usual finite intersection point. This is called *projective geometry*, and is described in more detail in the answer to another question.

In summary, then: in usual geometry, parallel lines do not meet. There is no such thing as infinity, and it is wrong to say that parallel lines meet at infinity.

However, you can construct other geometric systems, whose "points" include not only the points of familiar geometry (describable as coordinate pairs (x, y)), but also other objects. These other objects can be constructed in various ways, as described in the discussion of projective geometry. In these other geometric systems, parallel lines may meet at a "point at infinity". Whether this is one single point or different points for different classes of parallel lines, depends on the particular geometric system you are considering.

You may also be interested in our answers and explanations page, which contains a discussion of the question does infinity exist?

Constructing a Pentagon

Asked by Ting Ting Wu, student, State College Area High School on January 27, 1998:

> I want to know how to construct a pentagon. I have done it before, but I have forgotten how. I remember it is similar to constructing a hexagon, but a bit more difficult. Thanks.

There are several ways to do it. Unfortunately we are very short-staffed right now and cannot spare the resources to hunt down the easiest and most elegant construction. However, the following method will work:

Constructing a pentagon is equivalent to dividing a circle (a full 360 degrees) up into five equal parts (angle 72 degrees each). The cosine of 72 degrees is $(\sqrt{5} - 1)/4$ (this can be found by starting with the equation $\cos(5\theta) = \cos(360 \text{ degrees}) = 1$, using trigonometric identities to write $\cos(5\theta)$ as a polynomial in $\cos(\theta)$, factoring and solving the resulting polynomial equation for $\cos(\theta)$).

Therefore, this angle of 72 degrees can be constructed by building a right-angled triangle whose hypotenuse is 4 and whose adjacent side is of length $\sqrt{5}-1$. This latter length can be constructed by taking hypotenuse of a right triangle whose other sides have lengths 1 and 2, and subtracting length 1 from it.

The following procedure uses this idea to construct a pentagon:

Start with a circle C, with centre point O. Let P be a point on C. Draw the perpendicular bisector L to segment OP (bisecting it at point Q). Construct the midpoint R of OQ. (RQ is going to be our unit length).

With centre Q and radius RQ, draw an arc intersecting L at point S. Draw segment OS. (This is the hypotenuse of a right triangle OQS whose other sides have length 1 and 2, so OS has length $\sqrt{5}$).

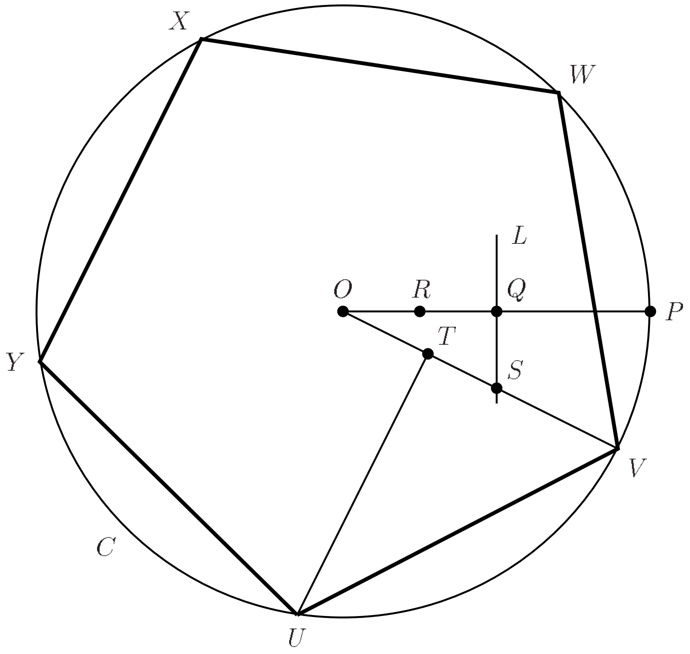
With centre S and radius RQ $(= QS)$, draw an arc intersecting OS at point T. (Now OT has length $\sqrt{5} - 1$).

Construct the line passing through point T at right angles to OT. Let it intersect the circle C at point U.

Now the triangle OTU has hypotenuse of length OU = radius of $C = 4$, and side length $OT = \sqrt{5}-1$. Therefore, angle UOT is 72 degrees. Extend segment OT past S until it meets the circle C at point V; you have now constructed two vertices (U and V) of the pentagon.

To construct the remaining vertices: with centre V and radius UV draw an arc intersecting C at point W. With centre W and the same radius, draw an arc intersecting C at point X. Finally, with centre X and the same radius, draw an arc intersecting C at point Y. $UVWXY$ will be a pentagon.

There are probably much more efficient ways to do it, but the above procedure will certainly work, for the reasons described. The procedure is illustrated below:



The Three Classical Impossible Constructions of Geometry

Asked by several students on August 14, 1997:

> I would like to know the three ancient impossible constructions problems using only a compass and a straight edge of Euclidean Geometry.

The three problems are:

1. *Trisecting an angle* (dividing a given angle into three equal angles),
2. *Squaring a circle* (constructing a square with the same area as a given circle), and
3. *Doubling a cube* (constructing a cube with twice the volume of a given cube).

People tried for centuries to find such constructions. It was not until the development of "abstract algebra" in the nineteenth century that it was proven these constructions were impossible.

For those of you who are interested, the basic idea of the proof is as follows.

Since the only shapes you can draw with a compass and straightedge are line segments and circles (or parts of circles), the only ways of constructing new points out of old points to to take the intersection of two lines, two circles, or a line and a circle. Now, if you write down the general equations for these intersections and try to solve for one of the coordinates of the intersection point, you will end up with either a linear or quadratic equation. The coefficients of the equation involve the coordinates of the old points (or various sums, differences, products, or quotients of them).

Linear equations can be solved by simple division: the equation $ax = b$ has as its solution $x = b/a$. Quadratic equations can be solved using the quadratic formula. In each case, we see that the only arithmetic operations required to calculate the new coordinate from the old coordinates are addition, subtraction, multiplication, division, and taking square roots.

Therefore, if you start with some initial points whose coordinates are all rational numbers, then apply any sequence of compass-and-straightedge construction techniques, the coordinates of the points you end up with will be a very special kind of number: they will be obtainable from the rational numbers by a sequence of operations involving only addition, subtraction, multiplication, division, and the extraction of square roots.

The reason the three classical constructions are impossible is that they are asking you to be able to construct points whose coordinates are not numbers of this type.

Proving that they are not numbers of this type requires some very advanced mathematics from an area called *Field Theory*. I'll sketch some of the essential ideas.

The key theorem is this:

> If x is a number obtainable from the rationals using only addition, subtraction, multiplication, division, and the taking of square roots, then x is a solution to some polynomial equation with rational coefficients. Moreover, if one factors out irrelevant factors from this equation until one gets down to an "irreducible" polynomial equation (one that can't be factored any further and still have rational coefficients), the degree of this polynomial will always be a power of 2.

Intuitively speaking: each time you take a square root, you usually double the degree of the polynomial required to represent the number. For example, $\sqrt{2}$ can be represented as a solution to the quadratic equation $x^2 - 2 = 0$, whose degree is 2. (Of course, it's also a solution to the cubic equation $x^3 - 2x = 0$ whose degree is 3, not a power of 2; but that's because this equation can be factored, and after factoring out the redundant factor of x you are left with the quadratic $x^2 - 2 = 0$, which is irreducible: it can't be factored any further and still have rational coefficients). If we take a square root a second time, getting the number $2^{1/4}$, it's a root of the fourth-degree polynomial $x^4 - 2 = 0$ (which is irreducible). A number like $\sqrt{2} + \sqrt{3}$, also obtained by twice taking a square root, is also the root of an irreducible fourth-degree polynomial: $x^4 - 10x^2 + 1 = 0$. A number obtained by taking a square root three times, like $2^{1/8}$, is typically the root of an irreducible 8th degree polynomial, like $x^8 - 2 = 0$.

This intuitive understanding is not always correct, though. For example, $\sqrt{2} + \sqrt{8}$, even though it is obtained using two square roots, is the same as $3\sqrt{2}$, which involves only one square root. So the intuitive argument only serves to show you that the claims of the theorem are *reasonable*. Actually proving the theorem (and proving, not just that there's *some* irreducible polynomial equation for x whose degree is a power of 2, but that *every* irreducible polynomial equation for x also has that same degree) involves many advanced ideas.

Using the theorem, it is easy to prove the impossibility of the three constructions:

Doubling a cube is impossible because if you start with a cube of side length 1, you would need to cunstruct a cube whose side length is the cubed root of 2. But the the cubed root of 2 is a solution to the irreducible equation $x^3 - 2 = 0$ whose degree, 3, is not a power of 2.

Squaring a circle is impossible because if you start with a circle of radius 1 you would need to construct a square whose side length is $\sqrt{\pi}$. But this is a so-called "transcendental number": it is not the solution to *any* polynomial equation with rational coefficients, let alone one whose degree is a power of 2.

Trisecting an angle is impossible because if you start with an angle of 60 degrees (which is easily constructible), you would then need to be able to construct an angle of 20 degrees. This would be equivalent to constructing a point whose coordinates are the cosine and sine of 20 degrees. This is impossible because cos(20degrees) is a solution to the irreducible polynomial equation

$8x^3 - 6x - 1 = 0$ whose degree, 3, is not a power of 2.

Using Geometric Postulates for Theorems in 3 Dimensions

Asked by Tim O'Brien, teacher, Bremen High School on September 20, 1997:

> I am trying to prove that any four noncoplanar points of a three space determine that three space, using the following postulates and theorems:
>
> P1: If a and b are distinct points, there is at least on line on both a and b.
>
> P2: If a and b are distinct point, there is not more than one line on both a and b.
>
> P3: If a, b and c are points not all on the same line, and d and e are distinct points such that b, c, and d are on a line and c, a, and e are on a line, there is a point f such that a, b, and f are on a line and also d, e, and f are on a line.
>
> P4: There exists one line.
>
> P5: There are at least 3 distinct points on every line.
>
> P6: Not all points are on the same line.
>
> P7: Not all points are on the same plane.
>
> Thm. 2-1 If two points of a line are on a given plane, then every point of the line is on that plane.
>
> Thm. 2-2 Any two distinct coplanar lines intersect in a unique point.
>
> Any help would be appreciated.
>
> Tim

An axiomatic system like this is not the usual method for studying geometry in three and higher dimensions. One would normally employ the language of vector spaces, linear independence, bases, and so on. That gives you a much cleaner theory that is dimension-independent.

To do it using postulates and theorems such as the ones you describe requires that you first of all give a definition of what a "three space" is! You could either do this axiomatically by giving postulates involving planes and three-spaces (similar to your existing postulates for points and lines), or else you could adopt some sort of definition (for example, saying that "the three-space generated by non-coplanar points A, B, C, and D is defined to be the set of all points P for which the (unique) line through P and D intersects the plane described by A, B, and C".

(This definition would not work for ordinary geometry since it would omit points that lie in the plane through D parallel to the plane through A, B, and C; however, your postulates do not describe ordinary geometry because of postulate P3. Your postulates are for a form of projective geometry, and the definition I gave above would work for that).

Of course, you would also need to define what you mean by a plane. In fact, you need to do that even before you can prove theorems 2-1 and 2-2; they do not follow from the postulates, since the postulates make no mention of planes except for P7 which is not enough to capture what "plane" means.

Then, the way you would prove your desired statement would depend on the particular definitions you chose for "plane" and "three-space". For instance, if you adopted the definition I suggested, your task would then be to prove

> If E, F, G, and H are in three-space(A, B, C, D) then three-space(E, F, G, H) = three-space(A, B, C, D).

You could start by proving basic facts about planes, such as the fact that if two points are in a plane then the entire line containing them is in the plane, and that any three non-collinear points in a plane determine that plane. Then you could try to prove these facts about three-spaces.

For example, to prove that if P and Q each belong to the three-space $S =$ three-space(A, B, C, D) and R is on the line through P and Q then R is also in S, you could start by saying that the line PD intersects plane ABC in some point P' and the line QD intersects the plane ABC in some point Q' (by the definition of three-space). Having already proven that when two points are in a plane the entire line joining them lies in the plane, it follows that the entire lines PD, QD, and PQ lie in the plane PQD, so all six points P, P', Q, Q', D, and R lie in the plane PQD. Thm 2-2 guarantees that the lines $P'Q'$ and RD intersect in a point X. Since P' and Q' are in plane ABC and X is on the line $P'Q'$, it follows that X is on plane ABC, so line RD intersects plane ABC, showing that R is in S.

Next you could show that if P, Q, and R are in S, then the entire plane PQR is in S. You could do this by letting T be some point in plane PQR, and observing that lines TP and QR intersect in some point X (Thm 2-2). By the previous result and the fact that X is on the line QR, we know X is in S; since T is on the line PX, it follows that T is in S also.

Now you're in a position to prove that every point P in three-space(E, F, G, H) is also in $S =$ three-space(A, B, C, D): line PH intersects plane EFG in a point X. Since E, F, and G are in S, so is X; since P is on line HX and both H and X are known to be in S, so is P.

Finally, you must prove the converse: that every point in S is also in three-space(E, F, G, H). You would use similar sorts of arguments to do this.

Please bear in mind, though, that if the particular definitions of "plane" and "three-space" you're working with are different from the ones I suggested, you would need different proofs, appropriate for whatever definitions or postulates you use for these concepts.

Counting Points and Lines Using Axioms

Asked by Bob Williams on January 9, 1998:

> Suppose I am given that
>
> - If P and Q are two points, there is exactly one line containing P and Q
> - If L is any line, there is a point P which does not lie on L
> - There are at least three points on every line
> - Any two distinct lines intersect at exactly one point
> - There exists at least one line.
>
> How do I prove that, if one line contains exactly n points, then
>
> 1. Every line contains exactly n points?
> 2. Every point lies on exactly n lines?
> 3. The space contains $n^2 - n + 1$ points and $n^2 - n + 1$ lines?

The following hints should help you answer this question.

First, try proving that, for any two lines L and M, there is at least one point R not on either of them. Do this using the fact that L and M each have at least three points, so you can find a point P on L which isn't the intersection point, and a point Q on M which isn't the intersection point. The line PQ will have a third point R on it (because every line has at least three points).

R is not on L (if it were, L and PQ would both be lines containing R and P. Since there is a unique line joining any two points, L would have to equal PQ, contradicting the fact that Q is not on L). Similarly, R is not on M.

Now you can prove that any two lines L and M have the same number of points. They have their intersection point X in common. For each remaining point P on L, the line RP intersects M in a point $f(P)$. Show that f establishes a 1-1 correspondence between the points on L (other than X) and the points on M (other than X).

To prove part (2), for any point P, show that there must be at least one line L not containing P. Every line through P intersects L in a point. Show that this process sets up a 1-1 correspondence between lines through P and points on L, proving that P is on exactly n lines.

Now pick a point P. Every other point must lie on a line through P. There are n such lines, with $n-1$ points on each, so there are $n(n-1)$ points not including P. Thus there are $n(n-1)+1$ points in total. A similar argument gives you the number of lines.

Deductive and Inductive Reasoning

Asked by a student at Winona Senior High School on January 28, 1998:

> I was talking with my geometry teacher the other day and we discussed inductive and deductive reasoning. He wanted me to find out exactly what they are and find an example just to see if I could do it. Can you help me answer this question?

"Deductive reasoning" refers to the process of concluding that something must be true because it is a special case of a general principle that is known to be true. For example, if you know the general principle that the sum of the angles in any triangle is always 180 degrees, and you have a particular triangle in mind, you can then conclude that the sum of the angles in your triangle is 180 degrees.

Deductive reasoning is logically valid and it is the fundamental method in which mathematical facts are shown to be true.

"Inductive reasoning" (not to be confused with "mathematical induction" or and "inductive proof", which is something quite different) is the process of reasoning that a general principle is true because the special cases you've seen are true. For example, if all the people you've ever met from a particular town have been very strange, you might then say "all the residents of this town are strange". That is inductive reasoning: constructing a general principle from special cases. It goes in the opposite direction from deductive reasoning.

Inductive reasoning is *not logically valid*. Just because all the people you happen to have met from a town were strange is no guarantee that all the people there are strange. Therefore, this form of reasoning has no part in a mathematical proof.

However, inductive reasoning does play a part in the *discovery* of mathematical truths. For example, the ancient geometers looked at triangles and noticed that their angle sums were all 180 degrees. After seeing that every triangle they tried to build, no matter what the shape, had an angle sum of 180 degrees, they would have come to the conclusion that this is something that is true of every triangle. Then they would have looked for a way to prove it using deductive reasoning; that is, deduce it as a consequence of other known general properties of triangles.

In summary, then: inductive reasoning is part of the discovery process whereby the observation of special cases leads one to suspect very strongly (though not know with absolute logical certainty) that some general principle is true. Deductive reasoning, on the other hand, is the method you would use to demonstrate with logical certainty that the principle is true.

Both are necessary parts of mathematical thinking. If you just started with the known properties of triangles and played around with them aimlessly using deductive reasoning, it is unlikely you would discover the fact that the angle sum is always 180 degrees (though if you did happen to discover it that way, you'd know it for certain). However, by noticing that it's true in all the examples you've ever seen, inductive reasoning leads you to suspect that this fact is true. Then, once your suspicions have given you a target and a direction for your deductive reasoning, you construct your rigorous logical proof using deductive reasoning.

The "inductive reasoning" mentioned above is nothing to do with the "principle of induction", which says that if you know something is true for the number 1, and if whenever it is true for one number it is also true for the next number, it is then true for every positive integer. Although this principle is a form of reasoning that gets you to a general principle from some individual cases (which is the reason for the name "induction"), it does so in a precise and logically valid way that is really a form of deductive reasoning if viewed in the correct way. When people refer to an "inductive proof", they generally mean a proof that uses the (logically valid) principle of induction, rather than meaning a form of (logically invalid) inductive reasoning in the sense described above.

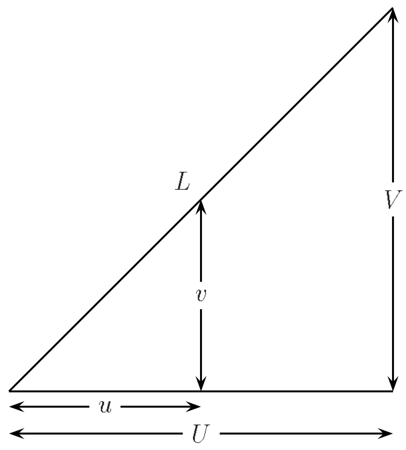
Questions about Slopes and Lines

Asked by Rafil Kroll-Zaidi, St. Stephen's School on October 11, 1996:

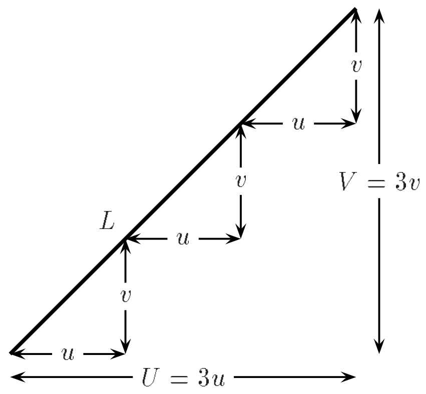
> I have several questions. First, why does $y = mx + b$ work?
>
> Secondly, how do you prove that something is an overdetermined system through graphing?
>
> Third, How do you prove that three lines which you can SEE are not collinear actually aren't?

1. I'm not entirely sure what you mean by "work" in your first question; let me assume that what you mean is "why can all lines (except for vertical lines) always be described by an equation of the form $y = mx + b$?"

   One elementary way to see it is to look at the picture below. If L is a non-vertical straight line, then the triangles below are similar triangles. The ratio V/U is the same as the ratio v/u. We call this ratio the *slope* m of the line.

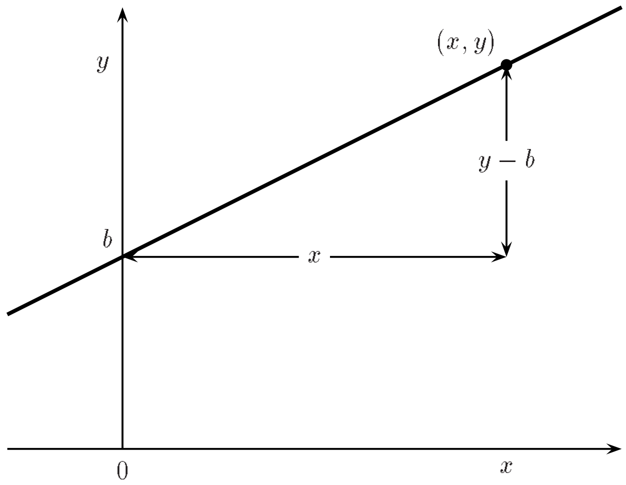


(For example, if U is three times u, then the picture below shows why V has to be three times v, so $V/U = (3v)/(3u) = v/u$. Similar reasoning can be used for the general case).



If we draw coordinate axes and let b be the value of y where the line passes through the y-axis, then we can ask ourselves, "what are the mathematical conditions that determine whether or not a point (x, y) lies on the line?"

Looking at the picture below, we see that $y - b$ is the "V" of the above picture and x is the "u", so the ratio $(y - b)/x$ equals m, in other words, $y - b = mx$, so $y = mx + b$.



That is why all lines (except for vertical lines, whose equation is of the form $x =$ constant) can be expressed in the form $y = mx + b$.

2. In general, graphing isn't a means of *proving* something is an overdetermined system, but rather a way get to get some visual insight into what the concept means.

   It really all depends on what you mean by "something". Are you talking in general, or are you referring to something specific like a set of linear equations?

   Without knowing the details of the context in which you're asking the question, let me just make some general comments.

   If you have a collection of equations, each of which has its solution set, then the set of points that satisfy *all* the equations is the intersection of these solution sets.

   For example, if you have the two equations $x + y = 1$ and $2x + y = 3$, the solution set to the first equation is the line $y = 1 - x$, and the solution set to the second equation is the line $y = 3 - 2x$. The set of points (x, y) that satisfy both the equations is the intersection of those two lines (which is the single point $(2, -1)$).

Now, if you start with just one equation, chances are the solution set has a whole range of points in it rather than just being a finite number of isolated points. For instance, the solution set to the equation $x + y = 1$ is an entire line.

However, as you add more equations, you are intersecting the solution set with other solution sets, so you get something smaller. A key question is, how many equations does it take until you get down to just a finite number of isolated points?

If your equations contain just two variables, then it generally takes two equations. The solution set to each equation is usually something like a line or a curve, and when you intersect them, you get a finite a number of points.

If your equations contain three variables, it generally takes three equations. The solution set to each equation will be something like a plane or other surface. Intersecting two of those gives you a curve, but intersecting that curve with a third surface will get you down to individual points. (For example, two adjacent faces of a cube will intersect along an entire edge, but three adjacent faces of a cube will intersect in a single point, namely a corner vertex).

And so on: if you have n equations, it generally takes n equations to reduce the solution set down to a single point or a collection of finitely many points.

(I say "generally" because you really need n *different* equations, and you may not always have this. For example, the two equations $y = x + 1$ and $2x = 2y - 2$ both have the same solution set, so specifying both of these equations together is no different from specifying just one of them. In this case, you'd need to add a third equation to get down to just a single point for the solution).

A collection of equations is *overdetermined* if it contains *more equations* than are needed to get the solution set down to a single point or finite collection of points.

For example, if you take the equations for three lines in a plane, they form an overdetermined system because just two lines are enough to get you a single intersection point. The third line might not even pass through that intersection point at all (in which case, there is no solution to the system of the three equations), or it might pass through it (in which case the third equation is redundant and could have been dropped altogether).

Note that an overdetermined system of equations usually has no solution. It is only in special cases that a solution exists. For example, the equations for three lines in a plane will have a common solution only if the lines are concurrent; generally, there will be no point that lies on all three lines.

By contrast, an underdetermined system usually does have a solution. It is only in special cases that no solution exists. For example, the equations for two lines in a plane will have a common solution in all cases except when the lines are parallel, for as long as the lines aren't parallel there will be an intersection point.

3. I assume you mean "cuncurrent" in your question (collinearity is a property of *points*, meaning they all lie on one line, concurrency is a property of lines, meaning they all pass through a single point).

   To prove three lines are not concurrent, you prove that there is no pair of numbers (x, y) which satisfies all three equations. One way to do that is to solve the first pair of equations and discover that there's only one solution (the point of intersection of the first two lines), then prove that this solution does not satisfy the third equation (i.e., that the intersection point of the first two lines does not lie on the third line).

For example, suppose you have the lines $y = x + 1$, $y = 2x + 3$, and $y = 3x + 2$. The intersection of the first two lines is given by $y = x + 1 = 2x + 3$, so $x + 1 = 2x + 3$, so $0 = x + 2$ (subtracting $x + 1$ from both sides), so $x = -2$, and $y = (-2) + 1 = -1$.

Therefore, $(-2, -1)$ is the point of intersection of the first two lines. But $(-2, -1)$ does not satisfy the equation of the third line, since -1 does not equal $3(-2) + 2$. This proves that the lines are not concurrent.

Why The Midpoint Formula Works

Asked by Nelson Siu, student, Gladstone Secondary on October 13, 1997:

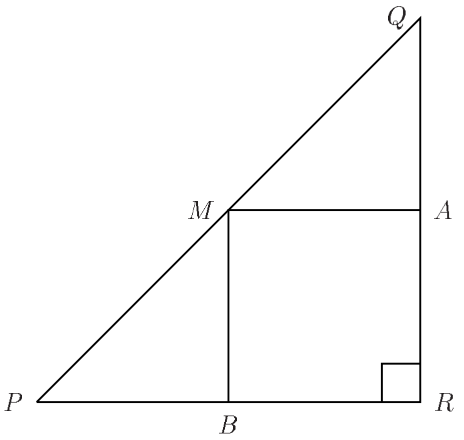
> I was asked to proof that the midpoint formula of a line works. Obviously it works, but how do I go about proofing that. I started off with a line PQ at a slope, and formed a right triangle. Next, the point M is b/n P and Q, and I guess I'm trying to show M is in the middle. After this, I'm kinda lost, so anything to get me back on track will be appreciated.
>
> Thanks in advance, I loved the work that you are doing here :-)

Suppose that $P = (x, y)$ and $Q = (X, Y)$ are the endpoints of our line segment. The midpoint M is then defined by $M = ((x + X)/2, (y + Y)/2)$. To show that M is really the midpoint of the line segment PQ, we need to show that the distance between M and Q is the same as the distance between M and P and that this distance is half the distance from P to Q.

One approach is to use the distance formula to calculate the relevant distances. The distance between P and Q is $\sqrt{(X - x)^2 + (Y - y)^2}$. The distances between M and P and M and Q are, respectively, $\sqrt{(\frac{x+X}{2} - x)^2 + (\frac{y+Y}{2} - y)^2}$ and $\sqrt{(\frac{x+X}{2} - X)^2 + (\frac{y+Y}{2} - Y)^2}$. Note that the distance from M to P simplifies to become $(\sqrt{(X/2 - x/2)^2 + (Y/2 - y/2)^2} = (1/2)\sqrt{(X - x)^2 + (Y - y)^2}$ which is just half the distance from Q to P. It is easy to see that the same holds true for the distance from M to P and therefore M is indeed the midpoint of the segment PQ.

Another approach is to make use of similar triangles. Let M be the midpoint of the line segment PQ (i.e. the point which is exactly half way between the two points). Now draw vertical and horizonal lines through all of the points. We will assume here that PQ is neither vertical nor horizontal, since it is simple to show that the midpoint formula is true for horizonal and vertical line segments.

The horizonal line though P and the vertical line through Q will meet at a unique point R. Similarly the horizontal line through M meets the vertical line through Q at a single point A and the horizontal line through P meets the vertical line through M at a single point B.



Note that PMB and MQA are similar triangles. Since $|PM| = |MQ|$ (because M is the midpoint of PQ), they are congruent triangles. Thus $|PB| = |MA| = |BR|$ so B is the midpoint of PR, and similarly A is the midpoint of QR.

The x-coordinate of M equals the x-coordinate of B. Assuming R is to the right of P as in the picture, this equals the x-coordinate of P plus $|PB|$, which is $x + |PB| = x + (1/2)|PR| = x + (1/2)(X - x) = (x + X)/2$. (If R is to the left of P, so is B, and we get the x-coordinate of M equal to $x - |PB| = x - (1/2)|PR| = x + (1/2)(x - X) = (x + X)/2$).

A similar argument shows the y-coordinate of M equals $(y + Y)/2$.

Infinity, Pi and Symmetry

Asked by Rudy Shearer, student, westbourne on July 3, 1997:

> Please tell me about the concepts of infinity, π and symmetry.

Before I begin, I will note that a truly detailed account of any of these topics could easily fill a book (I know that there is at least one book dedicated to the history of π and I have little doubt that there are entire books on symmetry as well). What follows are a few words about each topic.

For information about infinity, refer to the page Does Infinity Exist?.

Pi is, quite simply, a number. It has been known of since ancient times — there are references to it in the Bible and in Greek literature, to name a few. Over the years we have gained more and more knowledge of π and calculated it to progressively greater and greater precisions (it is now known to millions of decimal places).

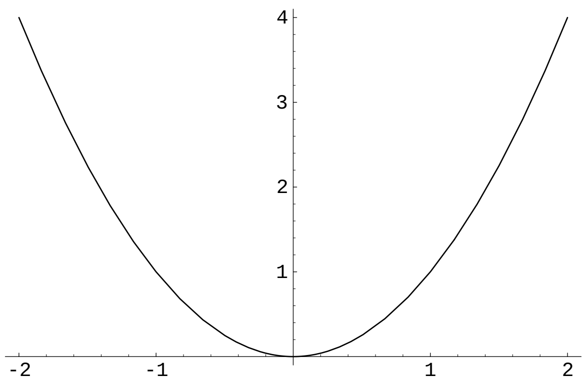
The definition of π is remarkably simple. It is the ratio between the circumference and diameter of a circle. To a few decimal places its value is $3.14159265\ldots$. Similar to the square root of 2, it is an irrational number and its digits never repeat. In this sense (and others) it is somewhat pathological.

Nevertheless, π turns up frequently in applied mathematics problems, sometimes in surprising ways. For instance, the sum $1 + 1/4 + 1/9 + 1/16 + 1/25 + \ldots$ converges to $\pi^2/6$ (you may see the proof of this in a calculus course some day).

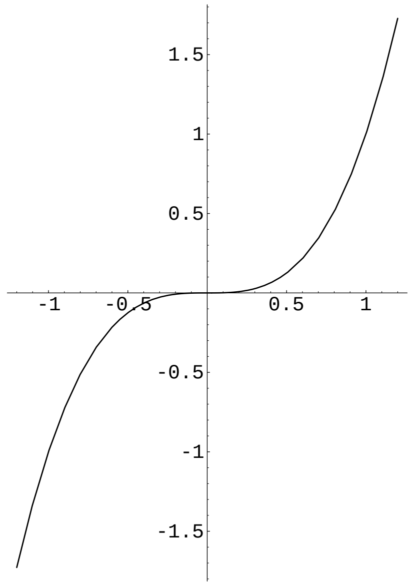
It turns out that there is a unit of angular measure frequently used in mathematics which involves π. If two lines intersect at a point, we can measure the angle between them as follows. Draw a circle (with any radius you like) centered at the point where the lines cross. Now the measure of the angle is given by the length of its arc on the circle divided by the radius of the circle. This is the measure of the angle in radians. There are 2π radians in a full circle and we can therefore convert degrees to radians by multiplying by $\pi/180$.

Symmetry is one of Mathematicians' greatest friends, as it often helps to simplify their lives. It is difficult to define exactly what it is, though it always refers to some sort of self-similarity of an object. This sort of repetition allows us to study a small part of an object and still have a good feel for how it behaves as a whole.

Often this symmetry is in the form of a reflection: one half of the object is the mirror image of the other half of the object. Consider the graph of the parabola $y = x^2$. The section of the parabola to the right of the y-axis is just a reflection of the other half of the parabola.



Other times symmetry occurs as reflection through a point. The graph of the curve $y = x^3$ exhibits this type of symmetry. If we wanted to find the signed area between this curve and the x-axis between the lines $x = -1$ and $x = +1$, we could note that the area to the left of the y-axis has the same value, with a change of sign, as the area to the right of the y-axis. Without performing any significant calculations we can use symmetry to deduce that the total area is 0.



Geometric solids often exhibit some sort of useful symmetry. Suppose we wanted to calculate the volume of an icosahedron (a 20 sided object whose faces are all equilateral triangles). The task seems difficult if we fail to notice that our solid is simply 20 identical pyramids glued together. We can just find the volume of one of these pyramids and then find the volume of the whole icosahedron by multiplying by 20.

Symmetry can occur in more subtle ways as well. Sometimes proofs themselves exhibit some symmetry. Instead of repeating one's self, authors often prove only one of the cases of a theorem. They then argue that the other similar cases are proved in a symmetric way.

Why is $e^{\pi i} = -1$?

Asked by Brad Peterson, student, Roy High on January 29, 1997:

> I was watching an episode of The Simpsons the other day, the one where Homer gets sucked into the third dimension, and in this 3-D world, there was an equation that said $e^{\pi i} = -1$. So I put it into the calculator and it worked, but I have no idea why, because e to any power isnt supposed to be a negative number, and I thought π was in no way related to e.
>
> If you could explain the process, it would save lots of time pondering and plugging e, π, and i into the calculator in random ways to figure out whats going on.

We'd be glad to explain; that's exactly what this area is here for.

The first question to ask, though, is not "why does $e^{\pi i} = -1$", but rather, "what does $e^{\pi i}$ even *mean*?" In other words, what does it mean to raise a number to an imaginary power?

Once that question is answered, it will be much more clear why $e^{\pi i} = -1$. It turns out that $e^{ix} = \cos x + i \sin x$ for all x, a fact which is known as *de Moivre's formula*, and illustrates how closely related the exponential function is to the trigonometric functions. From this formula, it follows immediately that $e^{\pi i} = \cos \pi + i \sin \pi = -1$.

So now, the question is, why is $e^{ix} = \cos x + i \sin x$ the "right" thing to define what e raised to an imaginary power means?

Raising a number to an imaginary power makes no sense based on the original definition of exponentiation you learned, where a^b means "a multiplied by itself b times." That definition only makes sense when b is a positive integer. After all, what would it mean to multiply something by itself i times??

Of course, the original definition doesn't even make sense for fractions and negative numbers. You should have learned how to extend the definition to include fractions. For example, since $1/3$ is that number which, when multiplied by 3, gives you 1, it makes sense to define $a^{1/3}$ to be that number which, if you raise it to the power of 3, would give you a^1 (i.e., a); in other words, $a^{1/3}$ is defined to be the cube root of a. Similarly, you learned how to extend the definition to negative exponents by $a^{-b} = 1/a^b$.

But none of these considerations give any clue as to what raising a number to a complex power should mean. Instead, we need to express exponentiation, or its properties, in some way that can be extended to complex powers.

The first way to do this is to use the fact that e^x happens to be equal to the infinite sum

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots$$

(where $n!$ means n factorial, the product of the numbers $1, 2, \ldots, n$).

The reason why this is so depends on the theory of Taylor series from calculus, which would take too long to describe here. You will encounter it in a calculus class at some point, if you haven't already.

Now, this infinite sum makes perfectly good sense even for imaginary numbers. By plugging in ix in place of x, you get

$$\begin{aligned} e^{ix} &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots \\ &= 1 + ix + i^2\frac{x^2}{2!} + i^3\frac{x^3}{3!} + i^4\frac{x^4}{4!} + i^5\frac{x^5}{5!} + \cdots \\ &= 1 + ix - \frac{x^2}{2!} - i\frac{x^3}{3!} + \frac{x^4}{4!} + i\frac{x^5}{5!} + \cdots \\ &= (1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots) + i(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots) \end{aligned}$$

Now it turns out that $1-x^2/2!+x^4/4!-\cdots$ is the infinite sum for $\cos x$, while $x-x^3/3!+x^5/5!-\cdots$ is the infinite sum for $\sin x$ (again by the theory of Taylor series). Therefore, $e^{ix} = \cos x + i \sin x$.

Now, this may be a little unsatisfying to you since I haven't explained why e^x, $\cos x$, and $\sin x$ equal those three different infinite sums. I can't do so without assuming some calculus background that you may not have.

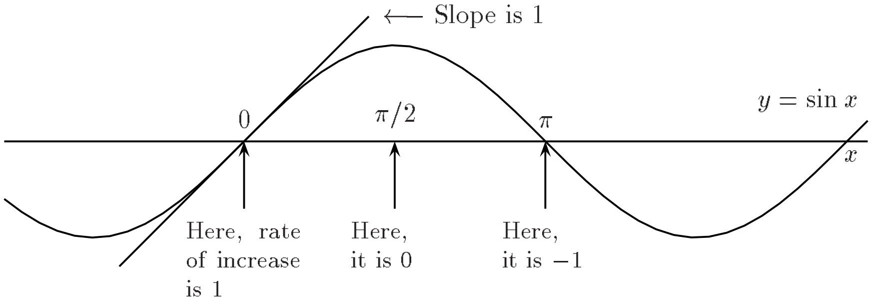
However, here's another way of understanding why $e^{ix} = \cos x + i \sin x$. It too involves some calculus, but I can describe the calculus involved more easily.

Associated to many functions $f(x)$ is another function $f'(x)$, called the *derivative* of $f(x)$. It measures how rapidly $f(x)$ is changing at the value x.

If $f(x) = e^x$, $f(x)$ may represent, for example, an exponentially growing population. The rate of change of such a population (the number of births per day, for example) is directly proportional to the current size of the population; that is, $f'(x)$ is a constant times $f(x)$. When $f(x) = e^x$, that constant is exactly 1 (that's the property which defines the number e). More generally, if $f(x) = e^{ax}$, then $f'(x) = ae^{ax}$.

What about the trigonometric functions? Well, if $f(x) = \sin x$, then $f'(x) = \cos x$, and if $f(x) = \cos x$, then $f'(x) = -\sin x$.

If you think about it for a minute, these equations are very reasonable. First of all, when $x = 0$, $\sin x$ equals zero but increases as x increases; in fact, the slope of the graph of $y = \sin x$ at the point $(0, 0)$ is 1, which is another way of saying that the rate of increase there is 1, so $f'(0) = 1$.



But then, as x increases to $\pi/2$, the rate of increase drops off and eventually $\sin x$ stops increasing altogether and starts decreasing. In other words, $f'(x)$ drops to zero when $x = \pi/2$, and becomes -1 by the time x reaches π (see the picture).

Therefore, $f'(x)$ is a function which starts at 1 when $x = 0$, decreases to 0 when $x = \pi/2$, drops to -1 when $x = \pi$, rises back to 0 when $x = 3\pi/2$, and so on. This is precisely what the cosine function does, so it should be no surprise that $f'(x) = \cos x$. Similar reasoning shows why it is reasonable that, when $f(x) = \cos x$, $f'(x) = -\sin x$. The exact proofs of these facts you will see in a calculus class.

Now, keeping those facts in mind, what should e^{ix} be? If we write it in terms of real and imaginary parts $g(x) + ih(x)$, what should the functions $g(x)$ and $h(x)$ be?

The key is to take the derivative. It is only reasonable to define e^{ix} in such a way that it still has the same properties as mentioned above, namely, the derivative of e^{ax} should still equal ae^{ax}. Therefore, if $f(x) = e^{ix}$, we should have

$$f'(x) = ie^{ix} = i(g(x) + ih(x)) = -h(x) + ig(x).$$

But $f'(x)$ should also equal $g'(x) + ih'(x)$, so we are looking for a pair of functions g and h for which $h' = g$ and $g' = -h$. This is exactly the same interrelationship that the sine and cosine functions have, as we saw above. It also turns out that these two equations, together with the conditions $g(0) = 1$ and $h(0) = 0$ that arise from the fact that e^{0i} needs to equal 1, uniquely determine the functions g and h.

It follows from all this that g must be the cosine function and h must be the sine function. That is why $e^{ix} = \cos x + i \sin x$.

What is the Square Root of i?

Asked by a student at Eastwood High on May 9, 1997:

> What is the square root of i?

There are two square roots of i: $(1/\sqrt{2})(1 + i)$ and $(-1/\sqrt{2})(1 + i)$. You can check that these are indeed square roots of i: just square each of them, and you get i.

The important question is, how are these answers obtained?

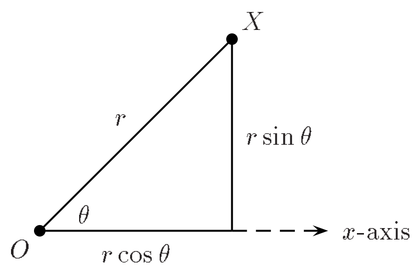
An elementary, but not the best, way to obtain this answer is to solve the equation $(a+bi)^2 = i$ for a and b. If you expand this equation using the rules for complex multiplication, you get $(a^2 - b^2) + (2ab)i = 0 + 1i$. Equating real and imaginary parts gives $a^2 - b^2 = 0$ and $2ab = 1$.

The equation $a^2 - b^2 = 0$ means $a = \pm b$. However, if you plug $a = -b$ into the second equation you get $-2b^2 = 1$ which can not be satisfied by any real number b. Therefore, the case $a = -b$ is not possible, meaning a must equal b. Then the second equation becomes $2a^2 = 1$. This means either $a = b = 1/\sqrt{2}$ or $a = b = -1/\sqrt{2}$. These are the answers that were given above.

However, the easiest and most insightful way to take the square root of a complex number (as well as any higher order roots) is to use the geometric representation of the complex numbers. Just as we can plot real numbers as points on a line, we can think of complex numbers as lying on a plane. The horizontal (real) axis corresponds to the real part of the complex number and the vertical (imaginary) axis corresponds to the imaginary part. For example, the number $3+2i$ is located at the point $(3, 2)$ which is two units above and three units to the right of the origin.

What is most useful about this visualisation of the complex numbers is how addition and multiplication behave. To add two numbers $a + bi$ and $c + di$, we can think of shifting the first one c units to the right and d units up.

Multiplication is a bit more difficult to see. Before we go on, it is useful to introduce another means of specifying points in the plane. Suppose that X is any point in a plane and suppose that O is the origin. If we know the distance r from X to the origin O and we also know the angle θ between the positive real axis and the line segment OX, we can locate the point X. This angle and distance are known as the polar coordinates of X. Multiplication corresponds to adding the angles of the two points and multiplying their lengths.



The reason for this has to do with trigonometry: the coordinates of the point whose distance is r and whose angle is θ are given by $x = r\cos\theta$, $y = r\sin\theta$. That means it is the complex number $r(\cos\theta + i\sin\theta)$ which equals $re^{i\theta}$ because of de Moivre's formula (which is explained in the answer to another question, the one on why $e^{\pi i} = -1$). If you multiply two numbers like that, you get

$$re^{i\theta}se^{iw} = (rs)e^{i(\theta+w)}$$

(the laws of exponents still hold for complex exponential). In other words, the product is a complex number whose distance from the origin is rs (the products of the distances of the factors), and whose angle is $\theta + w$ (the sum of the angles of the factors).

Understanding multiplication helps us understand other operations such as taking the square root. For the sake of simplicity, first consider the case in which the number we are interested in has a distance of 1 from the origin. The square root of this number also has a distance of 1 from the origin and forms an angle with the real axis which is 1/2 of the angle corresponding to the original number. In the event that the number is not 1 unit away from the origin, we obtain new distance from the origin by taking the square root of the old distance from the origin (here the lengths are positive real numbers and the notion of “square root” is already defined).

Now we can understand why we got the answer we did for the square root of i. It is easily seen that i forms an angle of 90 degrees with the real axis and has distance of 1 from the origin. Its square root is the number with a distance of 1 from the origin and an angle of 45 degrees from the real axis (which is $1(\cos 45° + i \sin 45°) = (1/\sqrt{2})(1 + i)$).

A cautious reader will note that there is some ambiguity in choice of the angle in our definition of polar coordinates. A point of distance 1 from the origin creating an angle of 45 degrees with the real axis is the same point which is 1 unit from the origin and forms an angle of 405 degrees with the real axis. Generally we always insist that the angle be between 0 and 360 degrees. Note however that when taking the square root of a complex number it is also important to consider these other representations. For instance, i can also be viewed as being 450 degrees from the origin. Using this angle we find that the number 1 unit away from the origin and 225 degrees from the real axis ($(-1/\sqrt{2})(1 + i)$) is also a square root of i.

What is i to the Power of i?

Asked by Oliver Varban, student, Earl Haig Secondary School on March 6, 1997:

> I am interested in knowing what i to the power of i is.

The first question to address is what it means to raise one complex number to the power of another. There is a basic definition of what it means to raise e to a complex power, as described in the answer to an earlier question. Therefore, if z is any complex number for which $e^z = i$, $(e^z)^i = e^{iz}$ is a possible value for i^i.

What are the possible values for z? Well, if we write $z = a + bi$, then $e^z = e^{a+bi} = e^a e^{bi}$. By de Moivre's theorem (explained in the answer to an earlier question), $e^{ib} = \cos(b) + i \sin(b)$, so $e^z = e^a(\cos(b) + i \sin(b))$. This expression equals i exactly when $a = 0$, $\cos(b) = 0$, and $\sin(b) = 1$. This occurs when $b = \pi/2 + 2n\pi$ for some integer n, so the possible values of z are $0 + (\pi/2 + 2n\pi)i$.

Therefore, the values of i^i are $e^{zi} = e^{(\frac{\pi}{2}+2n\pi)i \cdot i} = e^{-(\frac{\pi}{2}+2n\pi)}$

for any integer n.

Note that there is more than one value for i^i, just as 2 and -2 are both square roots of 4. (However, while the square roots of a number always have the same magnitude even if they differ in sign, the values of i^i have different magnitudes). The principal value of i^i would be $e^{-\pi/2}$—the case where $n = 0$.

It's also interesting to note that all these values of i^i are real numbers.

Raising a Number to a Complex Power

Asked by Wei-Nung Teng, student, Stella Matutina Girl's High School on June 17, 1997:

> How do you define a^{b+ci}? a, b, c are real numbers.
>
> example: $5^{3+2i} =$?
>
> Thanks

You can find a detailed answer to this question by reading the answers to some of the other questions on this site, but here is a summary of the answer all together in one place.

The ordinary definition of exponentiation of real numbers (a^x) only makes sense when x is rational. To extend the definition to irrational and then to complex values of x, you need to rewrite the definition in a way that makes sense even when r is complex.

One way to do this is to use the fact that e^x can be expressed as the infinite sum

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots$$

(where $n!$ means n factorial, the product of the numbers $1, 2, \ldots, n$).

It makes perfectly good sense to add and multiply complex numbers, and the theory about infinite sums can also be extended to complex numbers, so this formula can be used as a definition of what e^x means when x is complex.

If x is a "purely imaginary" number, that is, if $x = ci$ where c is real, the sum is very easy to evaluate, using the fact that $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, $i^5 = i$, etc. When you do this and split the sum into its real and imaginary parts, you find that the real part is the same as the infinite sum expression for $\cos c$, and the imaginary part is the same as the infinite sum expression for $\sin c$. This gives rise to *de Moivre's formula*:

$$e^{(}ic) = (\cos c) + i(\sin c)$$

This is explained in more detail in the answer to another question. Also given in that answer is an alternative explanation which doesn't rely on the fact that functions like e^x, $\cos x$, and $\sin x$ have representations as infinite sums. That's a fact you probably haven't seen yet, and it's hard to explain without some calculus background.

Now we know what e raised to an imaginary power is. One can also show that the definition of e^x for complex numbers x still satisfies the usual properties of exponents, so we can find e to the power of any complex number $b + ic$ as follows:

$$e^{b+ic} = (e^b)(e^{ic}) = (e^b)((\cos c) + i(\sin c))$$

Finally, for a real number a, you can define a^{b+ic} by writing $a = e^{\ln a}$:

$$\begin{aligned} a^{b+ic} &= e^{(\ln a)(b+ic)} = e^{(b \ln a)+i(c \ln a)} \\ &= e^{b \ln a}(\cos(c \ln a) + i \sin(c \ln a)) \\ &= a^b(\cos(c \ln a) + i \sin(c \ln a)). \end{aligned}$$

This answers the question you asked. Now, if a is a complex number instead of a real number, things are more complicated. There is no single value to "$\ln a$": there are lots of different complex numbers z for which $e^z = a$, and for any such complex number z, you could define a^{b+ic} to be $e^{z(b+ic)}$ and use the above technique to calculate it. This is illustrated in the answer to another question, where it is shown that the expression i^i has an infinite set of possible values.

In fact, the same thing is true even when a is a real number. Technically speaking, the expression a^{b+ic} has infinitely many possible values (except when b and c are both rational numbers), because instead of doing the calculation writing $a = e^{\ln a}$, you could also do it by writing $a = e^{\ln a + 2\pi i}$, or by writing $a = e^{\ln a + 4\pi i}$, or $a = e^{\ln a + 6\pi i}$, and so on. Each of these equalities is true (you can check them using de Moivre's formula to show that $e^{2\pi i} = 1$, so adding multiples of 2π to the exponent is the same as multiplying by 1, which doesn't affect the truth of the equality). Using these different equalities to do the calculation gives rise to infinitely many different possible values for a^{b+ci}.

When a is real it is more "natural" to use the ordinary real-valued logarithm $\ln a$ rather than than something like $\ln a + 2\pi i$. So it is reasonable to think of a^{b+ic} as having only one value (in much the same way as we think of $4^(1/2)$ as equalling 2 even though both 2 and -2 are square roots of 4). Technically, this value is called the *principal value*. This is what the formula up above gives you.

However, when a is not real there is no one natural choice of logarithm to prefer over any other, so in those cases we have to say that an expression like $a^(b+ic)$ has many different values.

The Origin of Complex Numbers and the Notation "i"

Asked by Brian (no last name given) on October 29, 1996:

> Who first thought up complex numbers?

Complex numbers were being used by mathematicians long before they were first properly defined, so it's difficult to trace the exact origin.

The first reference that I know of (but there may be earlier ones) is by Cardan in 1545, in the course of investigating roots of polynomials. During this period of time the notation $\sqrt{-1}$ was used, but more in the sense of a convenient fiction to categorize the properties of some polynomials, by describing how their roots would behave if we pretend that they have them. It was seen how the notation could lead to fallacies such as that described in the Classic Fallacies section of this web site, so $\sqrt{-1}$ was considered a useful piece of notation when putting polynomials into categories, but was not seen as a real mathematical object.

Later Euler in 1777 eliminated some of the problems by introducing the notation i and $-i$ for the two different square roots of -1. With him originated the notation $a + bi$ for complex numbers. He also began to explore the extension of functions like the exponential function to the case of complex-valued arguments. However, the numbers i and $-i$ were called "imaginary" (an unfortunate choice of terminology which has remained to this day), because their existence was still not clearly understood.

Wessel in 1797 and Gauss in 1799 used the geometric interpretation of complex numbers as points in a plane, which made them somewhat more concrete and less mysterious.

Finally, Hamilton in 1833 put complex numbers on a sound mathematical footing by showing that pairs of real numbers with an appropriately defined multiplication form a number system, and that Euler's previously mysterious "i" can simply be interpreted as one of these pairs of numbers. That was the point at which the modern formulation of complex numbers can be considered to have begun.

Asked by Megan Sullivan, student, Mater Dei High School on January 30, 1997:

> This is a question that has been bugging my A2T class - who first used i for imaginary numbers?

It was (as far as I know) the mathematician Leonhard Euler, in or around 1777.

For more information, see the answer to the question above.

Complex Numbers in Real Life

Asked by Domenico Tatone (teacher), Mayfield Secondary School on Friday May 3, 1996:

> I've been stumped!

> After teaching complex numbers, my students have asked me the obvious question: Where is this math used in real life!
>
> Your assistance would be greatly appreciated.

I'm going to give you my answer; if anybody else out there has some other particularly effective examples, please share them with us using the form below.

There are two distinct areas that I would want to address when discussing complex numbers in real life:

1. Real-life quantities that are naturally described by complex numbers rather than real numbers;
2. Real-life quantities which, though they're described by real numbers, are nevertheless best understood through the mathematics of complex numbers.

The problem is that most people are looking for examples of the first kind, which are fairly rare, whereas examples of the second kind occur all the time.

Here are some examples of the first kind that spring to mind. In electronics, the state of a circuit element is described by two real numbers (the voltage V across it and the current I flowing through it). A circuit element also may possess a capacitance C and an inductance L that (in simplistic terms) describe its tendency to resist changes in voltage and current respectively.

These are much better described by complex numbers. Rather than the circuit element's state having to be described by two different real numbers V and I, it can be described by a single complex number $z = V + iI$. Similarly, inductance and capacitance can be thought of as the real and imaginary parts of another single complex number $w = C + iL$. The laws of electricity can be expressed using complex addition and multiplication.

Another example is electromagnetism. Rather than trying to describe an electromagnetic field by two real quantities (electric field strength and magnetic field strength), it is best described as a single complex number, of which the electric and magnetic components are simply the real and imaginary parts.

What's a little bit lacking in these examples so far is why it is complex numbers (rather than just two-dimensional vectors) that are appropriate; i.e., what physical applications complex multiplication has. I'm not sure of the best way to do this without getting too far into the physics, but you could talk about a beam of light passing through a medium which both reduces the intensity and shifts the phase, and how that is simply multiplication by a single complex number.

Much more important is the second kind of application of complex numbers, and this is much harder to get across. I'm inclined to do this by analogy. Think of measuring two populations: Population A, 236 people, 48 of them children. Population B, 1234 people, 123 of them children. You might say that the fraction of children in population A is 48/236 while the fraction of children in population B is 123/1234, and that 48/236 (approx. 0.2) is much less than 123/1234 (approx. 0.1), so population A is a much younger population on the whole.

Now point out that you have used fractions, non-integer numbers, in a problem where they have no physical relevance. You can't measure populations in fractions; you can't have "half a person", for example. The kind of numbers that have direct relevance to measuring numbers of people are the natural numbers; fractions are just as alien to this context as the complex numbers are alien to most real-world measurements. And yet, despite this, allowing ourselves to move from the natural numbers to the larger set of rational numbers enabled us to deduce something about the real world situation, even though measurements in that particular real world situation only involve natural numbers.

In the same way, being willing to think about what happens in the larger set of complex numbers allows us to draw conclusions about real world situations even when actual measurements in that particular real world situation only involve the real numbers. You can point out that this happens all the time in engineering applications. If your students have seen some calculus, you can talk about trying to solve equations like $ay'' + by' + cy = 0$ (*) for the unknown function y. State that there's a way to get the solutions provided one can solve the quadratic equation $ar^2 + br + c = 0$ for the variable r. In the real numbers, there may not be any solutions. However, in the complex numbers there are, so one can find all complex-valued solutions to the equation (*), and then finally restrict oneself to those that are purely real-valued. The starting and ending points of the argument involve only real numbers, but one can't get from the start to the end without going through the complex numbers. Since equations like (*) need to be solved all the time in real-life applications such as engineering, complex numbers are needed.

Those are some thoughts on how I would try to answer the question "where are complex numbers used in real life".

Followup question by Greg Castle, Dickson College (Australia) on October 21, 1996:

> I am doing an assignment on complex numbers and their applications in the real world. I have found some fields where they are used, in engineering for example, but I really require formulas. Any formulas involving complex numbers that are used in the real world would be appreciated.

It's a little difficult to answer you're question without knowing what kind of formulas you're asking for (it's sort of like asking somebody for a sentence; you could write a sentence about just about anything!)

You can have formulas for simple laws; for example, the basic law relating current to voltage in a DC circuit, $V = IR$ where V = voltage, I = current, and R = resistance, generalizes through the use of complex numbers to an AC signal of frequency ω passing through a circuit with resitance, capacitance, and/or inductance, in the following way:

A sinusoidal voltage of frequency ω can be thought of as the real-valued part of a complex-valued exponential function

$$V(t) = Ae^{\frac{i\omega t}{2\pi}}.$$

Similarly, the corresponding current can be thought of as the real-valued part of a complex-valued function $I(t)$. These complex-valued functions are examples of the *second* kind of application of complex numbers I described above: they don't have direct physical relevance (only their real parts do), but they provide a better context in which to understand the physically relevant parts.

When such a voltage is passed through a circuit of resistance R, capacitance C, and inductance L, the circuit impedes the signal. The amount by which it impedes the signal is called the *impedance* and this is an example of the *first* kind of application of complex numbers I described above: a quantity with direct physical relevance that is described by a complex number. It is given by

$$Z = R + i\omega L + 1/(i\omega C)$$

and the circuit law becomes

$$V = IZ$$

where these are all complex numbers and the multiplication is complex multiplication.

So there's one example of a simple formula used in circuit analysis, generalizing the resistance-only case to the case of inductance, resistance, and capacitance in a single-frequency AC circuit.

Other formulas using complex numbers arise in doing calculations even in cases where everything involved is a real number. For example, there's an easy direct way to solve a first order linear differential equation of the form $y'(t) + ay(t) = h(t)$. But in applications, such as any kind of vibration analysis or wave motion analysis, one typically has a second order equation to solve.

Consider, for instance, the equation $y''(t) + y(t) = 1$. For a direct solution, one would like to "factor out" the differentiation and write the equation as $((d/dt)+r)((d/dt)+s)(y(t)) = 1$. Then you can let $g(t)$ denote $((d/dt) + s)(y(t))$, and we have the first-order equation $g'(t) + rg(t) = 1$ which can be solved for $g(t)$ using the method for first-order equations. Finally, you then use the fact that $y'(t) + sy(t) = g(t)$ to solve for $y(t)$ using first-order methods.

However, in order for $((d/dt) + r)((d/dt) + s)(y(t))$ to be the same as $y''(t) + y(t)$ (so that the method will work), it turns out that r and s have to be roots of the polynomial x^2+1, so we need $r = i$, $s = -i$. Therefore, passing through complex numbers gives a direct method of solving a differential equation, even though the equation itself and the final solution are all real-valued.

I hope the formulas in this and the previous example are of some use to you.

Asked by Melissa Bellin and Stephanie Carlson, students, Maple Grove Senior High on January 14, 1997:

> For my algebra class I need to find out how and why a specific job uses the square roots of negative numbers. Thanks for your help!

Some examples that come to mind are electrical engineers, electronic circuit designers, and also anyone in a profession where differential equations need to be solved. Besides, of course, mathematicians and physicists!

For more information, you might want to look at the answers given previously in this question.

More Complex Number Questions

Asked by Dr. Ethan B. Gallogly, UC Berkeley (PhD Chemistry) on October 14, 1996:

> Can you give one example of a directly measurable quantity which has an imaginary number in it? In all cases I've seen you either ignore the imaginary root or square it into a real quantity.
>
> Second, don't imarginary numbers arise because we ask that multiplication is communitive, i.e. $A \times B = B \times A$, but if we dropped this rule, and defined multiplication as $1 \times 1 = 1$, $-1 \times -1 = -1$, $-1 \times 1 = -1$, $1 \times -1 = 1$, then I believe we would lose the imaginary terms, but be stuck with a messier (similar to operator algebra) mathematics.

For your first question, here are some examples (see also the "answers and explanations" section of the web site, as well as an earlier question in the Question Corner area):

1. The strength of an electromagnetic field. This is a directly measurable quantity that is measured by a complex number. That number will be purely real if the field is all electric with no magnetic component, purely imaginary if the field is all magnetic with no electric component, and in other cases will have a non-zero real part and a non-zero imaginary part.

2. An attenuating medium placed in the path of an electromagnetic wave is also measured by a complex number. The number will be purely real if the medium affects the magnitude of the wave but leaves the phase unchanged, and purely imaginary if the medium shifts the

phase by 90 degrees. In general, a medium which scales the magnitude by a factor of R and shifts the phase by an angle t is described by the complex number Re^{it}.

A field F that enters medium M will emerge as F times M (where this product is by the laws of complex multiplication). It is this fact that justifies measuring the entire field by a single complex number rather than taking its electric and magnetic components separately and measuring each by a real number, for in the latter case you wouldn't have a way to multiply.

3. The state of a component in an electronic circuit is also measured by a complex number. That number will be purely real if there is a voltage across it but no current flowing through it (such as a fully charged capacitor) and purely imaginary if there is current flowing through it but no voltage across it (such as a fully discharged capacitor).
4. An LC filter in an electric circuit can be measured by a complex number, and the theory is much the same as for an attenuating medium placed in the path of an electromagnetic field.

As for your second question: commutativity is not the issue. More important is the fact that multiplication has to be related to addition by the distributive law.

There are plenty of non-commutative extensions of the reals (such as the quaternions), but any that satisfy the distributive law and are complete with respect to root-taking must of necessity contain the complex numbers as a subset.

Your example is simply a function from pairs of real numbers to real numbers for which there happens to exist a "root" r of -1 in the sense that $f(r,r) = -1$. There are plenty of such functions, but unless they are linear and distribute over addition, they don't provide the structure needed for a notion of "multiplication".

For instance, in your example you would have $1\times(1+(-1)) = 1\times 0 = 0$, but $(1\times 1)+(1\times(-1) = 1+1 = 2$.

Non-commutative multiplication is messier but still useful, but an operation that is not commutative, not associative, and doesn't distribute over addition cannot really play the part of multiplication in any reasonable mathematical theory.

Geometry and Imaginary Numbers

Asked by B. Delbecq on November 24, 1996:

> What kind of relationship can be made between geometry and Imaginary Numbers?

The set of complex numbers form a plane; that is, the complex number $a + bi$ corresponds to a point with coordinates (a, b).

Therefore, every complex number is a point on a plane. Equations that single out certain complex numbers over others correspond to various geometric figures. For example, the set of complex numbers whose magnitude is 1 forms a circle. The set of complex numbers whose imaginary part is 17 forms a line. And so on.

In higher dimensions, if you take equations involving several complex variables, the solution sets are geometric objects of various dimensions. For example, an equation such as $z^2 = w^3+1$ (where z and w are complex numbers) describes an interesting surface sitting inside 4-dimensional space. Understanding the geometric properties of surfaces like these, and their higher-dimensional analogues, is the aim of an important field of mathematics known as *Algebraic Geometry*.

You can also study the geometry of shapes given by equations involving purely real variables. However, it turns out that there is a far greater richness of structure in the complex case (where imaginary numbers are allowed), and many more important theorems that are true, than in the case of objects defined by equations involving real-only variables.

I don't know if this exactly answers what you were asking or not, but I hope it does.

Patterns in the Towers of Hanoi Solution

Asked by Alex Doskey on May 7, 1997:

> I first encountered the Towers of Hanoi puzzle when I was 8 years old. With an eager mind a attacked the puzzle and quickly discovered a pattern to its solution. This recursive solution is the one described in you web page discussion of this puzzle. At the same time I discovered another patten involved in moving the pieces, and since that time I have never seen any mention of this pattern in any discussion about this puzzle. This is not a difficult pattern, and I was wondering if I was just not reading the right articles.
>
> Although there is much more to it than this, here is the basic pattern that I discovered: Each piece in the puzzle moves in the same direction (clockwise or counterclockwise) throughout the entire solution of the puzzle. If you number the pieces (from smallest to largest or vice versa) all of the odd numbered pieces will move in one direction, and all the even numbered pieces will move in the opposite direction.
>
> The frequency of the moves is also predictable. The smallest piece will move every other turn, the next smallest every fourth turn, the next smallest every eighth turn and so on. These facts together lead to a very simple solution to the puzzle. Pick a direction to move the smallest piece. Move this piece every other turn. On the alternate turns, there is only one move available to you, so no thinking is required.
>
> This solution is much easier to work with (especially for larger stacks of disks), than the recursive solution, and yet I have never seen or heard any mention of it. My question is has anyone else ever seen this, or did I make an amazing discovery when I was a child?
>
> I do have more information about patterns involved in this puzzle if anyone has any questions. For example, how do you determine which direction to move the smallest disk if you want to move the stack from tower A to tower C and not tower B. The answer depends on if the stack has an odd or even number of disks. You can experiment with the ideas I presented in here, but I would recommend working with a physical representation of the puzzle, not a computer version. My first exposure was with towers and disks made of wood. If I hadn't had hands on experience with this puzzle, I may have never noticed these patterns.
>
> Alex

What you have described is actually the same solution as the recursive solution (that is, it is the same sequence of moves). However, you have described that sequence of moves in a way that makes it quite a bit easier to actually play the game. You are to be congratulated on noticing those patterns at such a young age, and more mention should be made of them in articles describing the game.

The reason there is a heavy emphasis made on the recursive nature of the solution, rather than on the patterns you have noticed, is that it is only through the recursive nature of the solution that fundamental theoretical questions can be answered. For example, how do you *know* that the pattern of moves you described will always solve the puzzle, regardless of the number of disks? And how do you know that it does so in the most efficient way? You can verify it by experiment, but experiment only carries you so far. For instance, you might verify it for 1 disk and for 2 disks and 3 disks and 4 disks and 5 disks, but then how do you know it's true for 6? And if you verify it experimentally for 6, how do you know it's true for 7 as well? And so on.

The recursive formulation allows you to easily answer those questions. For example, here's a proof that your pattern always works:

To move the pile of n disks from tower 1 to tower 3 (which we shall call the "counterclockwise" direction), you must move the top $n-1$ disks from 1 to 2 (which we shall call "clockwise"), then move the bottom disk from 1 to 3 (counterclockwise), then the top pile from 2 to 3 (clockwise).

Therefore, in moving a pile of n disks counterclockwise, the bottom disk will move counterclockwise and the remaining $n-1$ disk pile will move clockwise.

Therefore, the bottom disk of that remaining $n-1$ disk pile (which is the second from the bottom in the original pile) will move clockwise, and the remaining $n-2$ disk pile will move counterclockwise.

Therefore, the bottom disk of that remaining $n-2$ disk pile (the third from the bottom in the original pile) will move counterclockwise, and the remaining $n-3$ disk pile will move clockwise.

And so on. To make the argument mathematically rigorous, you would use the principle of induction, but I want to keep this answer at an informal level. You end up with the fact that the bottom, third from the bottom, fifth from the bottom, etc. must all move counterclockwise, while the remaining disks must move clockwise. This is exactly the same pattern you noticed, except now it is no longer in the category of something that has been observed to be true in all the cases one has happened to verify, but it is in the category of something which is known to be true with full mathematical certainty, because it can be logically proven.

You can derive the rest of your patterns, about which disks move when, in a similar way.

The bottom line is essentially this: the patterns you noticed are probably the easiest way to play the game, and the easiest way to physically accomplish the task of moving the disks. However, the confidence that these patterns actually work, that they actually form the minimal way of moving the disks, comes from the recursive formulation of the solution and the powerful mathematical arguments that can be derived from it.

Together, these two things illustrate the two important halves of mathematical discovery: the experimentation and noticing of patterns (such as you noticing the clockwise-counterclockwise method of solving the Towers of Hanoi puzzle), coupled with the formulation of them in a way which leads to the discovery of mathematical arguments that allow one to know with complete certainty why the patterns are the way they are (such as the recursive formulation of the Towers of Hanoi solution).

Musical Frequencies and Other Questions

Asked by wayneb@netnitco.net on August 4, 1996:

> In a sphere...why do we ignore the other 3/4 of the sphere and concentrate only on the positive and negative's of the X, Y, and Z axis ? Are the other areas less important? If so Why?
>
> #2 In the mathmatical progression on musical harmonics why do we only busy ourselves with the pleasant tones. Do these others not pose possibilities to door-

ways not yet opened?

#3 Is mathmatics so sacred that to question its validity is to bring upon oneself damnation and excomunication? Am I missing something here?

I'm going to answer your second question first. The question of what is and isn't a "pleasant" tone, and what possibilities the others offer, is really not a mathematical question at all. It comes down to a matter of physics. I'll try to briefly indicate the physics involved.

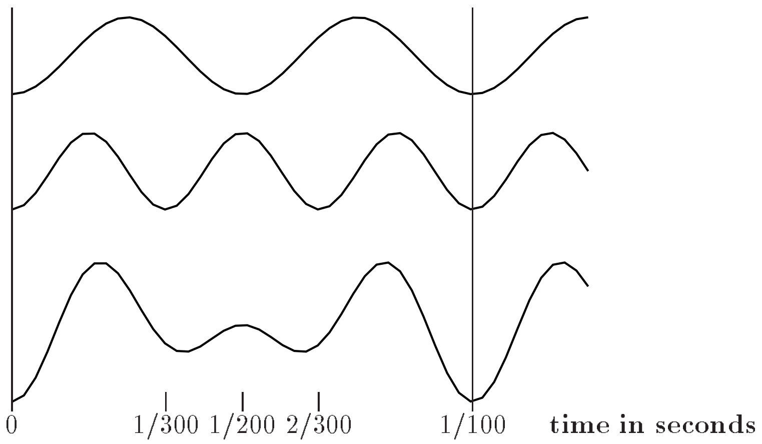
Suppose you start with one tone at a certain frequency, and then sound another one along with it. What will the result be? Well, if we think of both tones starting off together at the beginning of their respective pulses, a certain amount of time will elapse before they get back "into synch" with each other again, and during that time there will be an ever-changing sound.

Consider two examples:

1. A frequency of 200 beats per second, and one of 300 beats per second.

   The second tone will complete three full cycles in the same amount of time (1/100th of a second) that the first one takes to complete two cycles, so there is a changing sound-pattern that lasts for 1/100th of a second, which then repeats itself over and over.

   I've illustrated this in the picture below. The top shows the first wave, the middle shows the second wave, and the bottom shows the combination (sum) of the two.



   Although the first wave repeats itself after 1/200th of a second, and the second wave repeats itself after 1/300th of a second and again after 2/300th of a second, it's not until 1/100th of a second that the combined pattern starts repeating itself.

2. For our second example, Suppose that one had a tone of 200 beats per second and one of 303.91 beats per second. If you draw a diagram similar to the above, you'll find that it takes a full 100 seconds (20,000 beats of the first tone and 30,391 beats of the second tone) before the pattern starts repeating itself again.

These have quite a different effect on the ear. In the first case, the combined tone consists of a 1/100th-of-a-second pattern being constantly repeated. The fact that the 1/100th-second fluctuation has a complicated shape (rising all the way up, falling partway, rising a little, falling again, rising all the way, then finally falling all the way) gives the tone richness. The fact that each fluctuation lasts only 1/100th of a second means that it isn't perceived by the ear as causing the sound to change over time; it sounds the same from beginning to end. If you sound the two tones together for 10 seconds, the sound at the end of those 10 seconds will be the same as the sound at the beginning.

However, in the second case, you have an ever-changing sound that lasts 100 seconds before repeating! If you sound the two tones together for 10 seconds, the sound at the end of those 10 seconds may be quite different from the sound at the beginning.

For musical purposes, the first kind of sound has been traditionally judged "pleasing" to the ear, because it's a tone that's rich in texture and always sounds the same. The second will be full of discernible "beats" and other long-term modifications to the sound, that most people judge "ugly".

Mathematics comes into play because, to figure out the length of time it takes before the pattern repeats itself, you need to find when, if both tones start out beginning a pulse together, they next get back in phase with each other (i.e., when they next begin a pulse together). When that happens, tone 1 will have completed some whole number "a" of pulses, and tone 2 will have completed some whole number "b" of pulses, so you'd have

$$(a)(\text{time taken by one pulse of tone 1}) = (b)(\text{time taken by one pulse of tone 2})$$

i.e.

$$a/(\text{frequency 1}) = b/(\text{frequency 2})$$

i.e.

$$(\text{frequency 1})/(\text{frequency 2}) = a/b.$$

Thus, if the ratio of the two frequencies is expressible as the ratio of two small whole numbers (as our first example was; the frequencies were in a 3/2 ratio), the combined tone will be a short pattern repeated often. But the larger the numbers a and b have to be, the longer the pattern is.

It is when a and b are small numbers like 1, 2, 3, and 4 that you get the "purest" combined tones with the shortest fluctuations. That is why the standard musical scale is based on intervals such as a 2:1 frequency ratio (one "ocatve"), a 3:2 frequency ratio (one "fifth") and a 4:3 frequency ratio (one "fourth"), along with a 2:1 frequency ratio (one "octave"). One can employ mathematical analysis to figure out other notes of the scale from these.

The standard scale is obtained by starting with one tone and continually taking 3/2 ratios until one gets back to something very close to an exact number of octaves above the original tone. It takes 12 times before you get back, and this gives the standard 12-note system. (One doesn't get back exactly, but one gets back close enough that the ear doesn't notice the difference).

There is no mathematical reason at all why one cannot employ other scales. As I said, it's not a mathematical question what scale one wants to use. The standard one just happens to be the one that involves the most intervals with the "purest" tone (shortest length of the fluctuations in the tone). The same kind of mathematical analysis would apply to any musical system one wanted to consider. The study of the musical tones of our standard scale just happens to be one particular application of the mathematical theory of numbers; it is not itself a part of mathematics. The same mathematics would apply equally well to any other musical system.

Now, for your first question: I really don't know what you're referring to. First of all, there's no natural way to partition a sphere into quarters. You can, however, partition a sphere centred a the origin into *eighths*, called *octants*, based on the signs of the coordinates x, y, and z; perhaps that is what you are referring to.

If so, then the so-called "first octant" (where x, y, and z are all positive) is no more special than anywhere else on the sphere. The only reason one might consider it separately is if one is applying some technique where the sign of x, y, or z makes a difference in the precise details of the technique. This happens sometimes in multivariable calculus. In cases like this, one often gives the argument for the first-octant case and then, rather than writing it all out again for the other seven octants, leaves it as an exercise to the reader to make the appropriate changes to do the other cases.

Also, if one wants to give a problem involving, say, part of a sphere, to test a student's understanding of how to handle surfaces with boundaries, one would tend to choose the first-octant part so that the student can focus on the problem at hand and not have the main idea obscured by extra fiddly little details like putting all the minus signs in the right place.

But, mathematically, the first octant is no more special than any other part of a sphere.

As to your third question: mathematics is concerned with what is and isn't true. The entire basis of mathematics is asking questions, and coming up with answers that decide the question beyond any possibility of doubt through strict, logical argument. So there's nothing at all wrong with the concept of "questioning" in mathematics; it's what the subject is all about.

However, it's also important to realize that mathematics is a matter of fact and proof, not of opinion; one cannot argue against a proven fact, unless the proof is incorrect and one can demonstrate where the incorrectness lies. It's also important to realize that a lot of what masquerades as "mathematics" really is not, but rather is a combination of mathematics together with some assumptions about some real-life things. Perfectly valid mathematics can sometimes be used to draw incorrect conclusions about real-life situations, through error in what one was assuming about the situation. Questioning those conclusions is not at all the same thing as questioning the validity of the mathematics.

The Mathematics of Drum Design

Asked by Frank Gilbert, student on November 16, 1996:

> Howdy folks!
>
> I have a question for all you math wizards out there! I make african drums called Ashiko's as a hobby. Until now I have just been guessing the bottom hole size and coming up with some pretty good sounds...but I know I could perfect the tone quality if I just knew how to calculate the correct hole size in relation to the top hole and the length. It would sorta be like calculating the port size for a speaker box in a system with a totally enclosed speaker.
>
> For instance. One very popular drum I make is a 12" Ashiko. It has a 12" diameter top hole and a 6" diameter bottom hole with a length of 24". The Ashiko is shaped like a hiway pylon. I stretch a thin, goatskin head over the top hole and lash it tight with 4mm mountain climbing accessory rope.
>
> Any ideas?
>
> River Man

That's not an easy question to answer! Mathematically modelling the flow of sound in an enclosed 3-dimensional space requires some advanced techniques in the field of "partial differential equations", and many of the solutions require numerical approximation techniques to actually come up with numerical answers, since many of the functions involved can't be expressed in terms of familiar, everyday functions like addition, multiplication, exponentiation, trigometric functions, etc.

The 1-dimensional situation (e.g., a guitar string, or a thin pipe) is quite easy to analyze. Think about a wave of sound stretched out along the string or pipe, starting with zero amplitude at the end, and rising and falling as you move along the pipe. The amplitude has to be back at zero at the other end (for instance, in a vibrating string, the ends are tied down and not free to move).

Over the course of one cycle (one wavelength), the displacement of the string (or the compression of air in a pipe) starts at zero, rises to a positive value, drops back down through zero to a negative value, then rises again. So the only places at which it is zero are at the start and end of cycles, and half-way through. In other words, it is zero at the starting end of the pipe or string, and at a half-wavelength distance from the end, and at a full wavelength distance from the end, and at a three-halves-wavelength distance, and so on: at distances which are an integral multiple of $\lambda/2$ where λ is the wavelength.

In order for sound to resonate in the string or pipe, the displacement must be zero at the finishing end as well as the starting end, so the length L of the pipe must be an integral multiple of $\lambda/2$. Thus, the only wavelengths that will resonate are when $\lambda = 2L$, $\lambda = \frac{2L}{2}$, $\lambda = \frac{2L}{3}$, etc.

Frequency ω is related to the wavelength by $\omega = c/\lambda$ where c is the speed of sound. Therefore, the resonant frequencies are $c/(2L)$, $(2c)/(2L)$, $(3c)/(2L)$, and so on; in other words, the fundamental frequency of the string or pipe is $c/(2L)$; the others are higher octaves. The speed c of sound in a pipe depends on the air density, humidity, temperature, altitude, etc. In a string it depends on the string material and, most importantly, on the string's tension; that's why changing the tension changes the frequency of sound produced.

However, none of this simple analysis applies to your case, where you are dealing with sound waves in three dimensions. I do not know of any easy answer to your question. I will, however, make sure that your question is seen by others at the University; someone who works in applied mathematics and partial differential equations, or in acoustically related engineering, may be able to offer some insight.

Probably the best thing for you to do is to hunt down a book on the construction of such drums; it would likely contain the measurements for an optimal sound. However, you should realize that those measurements (while as accurate in practice as any computer's computation) were likely obtained by good old fashioned trial and error! In truth, these sorts of measurements are always *better* than those a computer gives you because there are always many overlooked discrepancies between the mathematical model of the drum and the actual drum (every piece of wood is different, as is every piece of goatskin).

While the modelling process is fascinating in its own right, trial and error (or even better someone else's trial and error) may well be the best route in this particular case.

Multiplying Matrices

Asked by David Dymov, student, P.C.V.S. on May 27, 1997:

> How do you multiply 2 matrices which have 4 numbers each?

It is perhaps just as easy to answer the much more general question of how two matrices should be multiplied together.

Suppose that A and B are two matrices and that A is an $m \times n$ matrix (m rows and n columns) and that B is a $p \times q$ matrix. In order for us to be able to multiply A and B together, A must have the same number of columns as B has rows (ie. $n = p$). The product will be a matrix with m rows and q columns. To find the entry in row r and column c of the new matrix we take the "dot product" of row r of matrix A and column c of matrix B (pair up the elements of row r with column c, multiply these pairs together individually, and then add their products).

For example suppose we have the matrices

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 9 & 10 \\ 11 & 12 \\ 13 & 14 \end{bmatrix}.$$

Their product is

$$AB = \begin{bmatrix} (1 \times 9) + (2 \times 11) + (3 \times 13) & (1 \times 10) + (2 \times 12) + (3 \times 14) \\ (4 \times 9) + (5 \times 11) + (6 \times 13) & (4 \times 10) + (5 \times 12) + (6 \times 14) \end{bmatrix} = \begin{bmatrix} 70 & 76 \\ 169 & 184 \end{bmatrix}$$

For those of you who would like an explict formula, here it is:

$$C_{r,c} = \sum_{i=1}^{n} A_{r,i} \, B_{i,c}$$

(where $C_{r,c}$ is the entry in row r and column c of the product matrix $C = AB$).

Why is multiplication of matrices defined in this complicated way? It is because matrices can be interpreted as ways of transforming one set of values into another set of values, and matrix multiplication corresponds to doing one transformation after another.

For example, matrix A, with its two rows and three columns, might describe a chemical reaction that starts with two types of input chemicals (let's call them X and Y) and produces three types of output chemicals (let's call them P, Q, and R). The numbers in the first row describe how much of each output chemical is produced from one unit of the first input chemical (X). The numbers in the second row describe how much of each output chemical is produced from one unit of the second input chemical (Y). That is, the numbers in

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

mean that each unit of chemical X produces 1 unit of P, 2 units of Q, and 3 units of R, while each unit of chemical Y produces 4 units of P, 5 units of Q, and 6 units of R.

Matrix B, with its three rows and two columns, could describe another chemical reaction that transforms the three chemicals P, Q, and R into two other chemicals, U and V.

What is the result of performing both chemical reactions, A and then B? You start with chemicals X and Y, then eventually end up with chemicals U and V. The matrix product AB describes how much of each output chemical you end up with.

For example, if you start with one unit of chemical X, it will under reaction A turn into 1 of P, 2 of Q, 3 of R. Under reaction B each unit of P will turn into 9 of U plus 10 of V; each of the 2 units of Q will turn into 11 of U and 12 of V; and each of the 3 units of R will turn into 13 of U and 14 of V. The total number of units of U we will end up with is therefore $(1)(9) + (2)(11) + (3)(13)$, and the total number of units of V is $(1)(10) + (2)(12) + (3)(14)$. These are the two numbers in the top row of the matrix product AB. The bottom-row numbers tell how many units of U and V you end up with if you start with one unit of Y (instead of starting with 1 unit of X).

It is this interpretation of matrix multiplication as the combination of two transformations that leads to the way matrix multiplication is defined.

How To Graph The Inverse Of A Function

Asked by mary jones, student, Dan River High School on September 27, 1997:

> How do you graph an inverse function once you have solved it?

It is actually easier to graph the inverse of a function than it is to solve for it.

First let's think about what it means to graph a function. Suppose that we wanted to graph a function $f(x)$. To do this, we would substitute numerical values for x and plot those ordered pairs (x, y) for which $y = f(x)$. For instance, if $f(x) = x^2$, we might try plotting the points $(0, 0)$, $(1, 1)$, $(2, 4)$, $(3, 9)$, etc.

Now how would we plot the inverse of a function? If g is the inverse function of f, then to graph g we'd plot the points (x, y) where $y = g(x)$. This condition is the same as the condition $x = f(y)$, so the graph of the inverse consists of points $(f(y), y)$.

These are the same as the points on the graph of f but with the order of the coordinates interchanged: instead of plotting $(x, f(x))$ for various numerical values of x, we plot $(f(y), y)$ for various numerical values of y (which is the same thing as plotting $(f(x), x)$ for various numerical values of x). For example, some of the points on the graph of the inverse of $f(x) = x^2$ are $(0, 0)$, $(1, 1)$, $(4, 2)$, and $(9, 3)$. Geometrically, we have just reflected the graph of the function $f(x)$ through the line $y = x$ to get the graph of the inverse of $f(x)$.

This method of graphing the "inverse" of a function always works, even when the function doesn't have an inverse. If the function doesn't have an inverse, it is because there are two distinct values a and b which we can assign to x to get the same value for $f(x)$. If we examine our function $f(x) = x^2$ we will note that $f(2) = f(-2) = 4$. The corresponding points on the graph of our "inverse function" are $(4, 2)$ and $(4, -2)$. Thus the graph which we constructed in this method is not really the graph of a function, since the value of the inverse of $f(x)$ is not well defined at 4 (it could either be 2 or -2).

Even though this approach will not always give us the graph of a function, it will whenever the inverse of $f(x)$ exists.

Does Every Function Have an Antiderivative?

Asked by Sam Beroz on Monday Dec 11, 1995:

> Does there exist a function for every antiderivative?

I'm interpreting your question to mean "does every function have an antiderivative?", which isn't quite what you wrote, but I think it's what you meant.

For continuous functions, the answer is yes. If you start with any continuous function $f(x)$ and want to find an antiderivative for it, you can look at the definite integral

$$F(x) = \int_0^x f(t)dt.$$

One form of the fundamental theorem of calculus says that derivative of this is $f(x)$. ($F(x)$ is the area under under the graph of f and above the interval from 0 to x. If you ask what is the rate of change of this area as x increases, the answer is exactly the height of the graph at x, that is, $f(x)$).

So, $F(x)$ is an antiderivative of $f(x)$. And, the theory of definite integrals guarantees that $F(x)$ exists and is differentiable, as long as f is continuous.

Most functions you normally encounter are either continuous, or else continuous everywhere except at a finite collection of points. For any such function, an antiderivative always exists except possibly at the points of discontinuity.

For more exotic functions without these kinds of continuity properties, it is often very difficult to tell whether or not an antiderivative exists. But such functions don't normally arise in practice.

Follow-up question from Sam Beroz on Wednesday Dec 13, 1995:

> How about when you are given a derivative of an unknown function and asked to intergrate it. Is there always an answer no matter how complex? Or are there some equations which may never serve as derivatives?

There is always an answer (there is always a function whose derivative is the function given to you, provided it is continuous).

However, it may not be possible to express the answer in terms of familiar functions and operations. For example, the antiderivative of e^{x^2} exists, but there is no simpler way to write the function other than to simply say "the antiderivative of e^{x^2}". You can't find a formula for it in terms of familiar functions, but it exists nonetheless.

Follow-up question by Hossien Esmaeili on December 4, 1996:

> Is there any better way of expressing the antiderivative of e^{x^2}?

No; it cannot be expressed any more simply than that.

In some situations (especially in probability) one needs to work with the antiderivative of e^{-x^2} a lot. Actually, in practical examples what comes up a lot is the antiderivative of $\frac{2}{\sqrt{\pi}}e^{-x^2}$. Because it occurs so frequently, it has been given a name, called the "error function", denoted $\operatorname{erf}(x)$.

Using complex numbers, you can express the antiderivative of e^{x^2} in terms of this "erf" function: it is $\frac{\sqrt{\pi}}{2i}\operatorname{erf}(ix)$.

But all this is doing is relating the antiderivative of e^{x^2} to another antiderivative that also cannot be expressed any more simply, but which comes up frequently enough that people have given a special name to it.

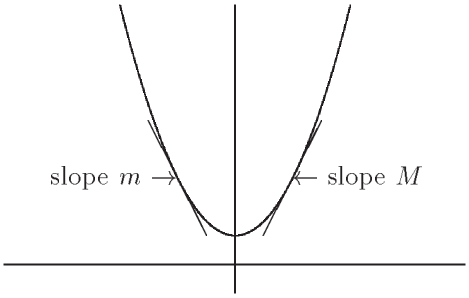
Symmetry of Functions and their Derivatives

Asked by Sam Beroz on Monday Dec 11, 1995:

> For every fuction with symmetry will the first derivative have symmetry of the other type? What Theorem proves this if it is in fact true?

Yes, it is true. If f is an even function (that is, has the same value if you replace x by $-x$), then its derivative will be an odd function (changes sign when you replace x by $-x$), and vice versa.

This is quite clear geometrically; in the picture below, for example, it is apparent that the slopes m and M are negatives of each other. You could even turn this into a geometric proof: if f is even, its graph is the same if you reflect it in a mirror placed along the y-axis, and therefore the tangent line at one point is the mirror reflection of the tangent line at the reflected point, and a line reflected in the y-axis has its slope multiplied by -1.



In the above picture, $m = -M$.

The way that you prove it using only calculus theorems (without needing any geometry at all) is as follows.

If f is an even function, that means that $f(x) = f(-x)$.

Now differentiate both sides. The left-hand side becomes $f'(x)$, and the right-hand side becomes $-f'(-x)$ (using the chain rule).

Therefore, $f'(x) = -f'(-x)$. In other words, the value of f' at x is the negative of its value at $-x$, so f' is an odd function.

Similarly, if you started with an odd function f, you have $f(x) = -f(-x)$. Differentiating both sides gives $f'(x) = +f'(-x)$, so f' is an even function.

In either case, f' has the opposite type of symmetry (even or odd) from f.

Factorials of Non-Integral Values

Asked by Richard Gillion on January 19, 1998:

> Someone said that the "factorial" of 0.5 is $\sqrt{\pi/4}$. It is indeed the only value which gives a smooth curve for "factorials" 1.5, 2.5, 3.5, ... continuous with the more familiar integer values. What possible connection is there between factorials and pi?

The "factorial" concept applies only to non-negative integers, so strictly speaking there is no such thing as the factorial of 0.5. However, there is an important mathematical function called the *Gamma function*, defined by

$$\Gamma(s) = \int_0^\infty t^{s-1} e^{-t} \, dt,$$

and if you define $f(x) = \Gamma(x+1)$ (the integral of $t^x e^{-t} dt$), then $f(x)$ has the same fundamental property that factorials have: $f(x) = xf(x-1)$ for all $x > 0$. (The reason is given below). Using this, and the fact that when you calculate $f(0)$ you get 1 which is the same as 0!, you can prove by induction that $f(x) = x!$ when x is a non-negative integer.

Therefore, this function $f(x)$ provides a natural extension of the factorial concept to all non-negative real numbers. What is meant by "the factorial of 0.5" is really

$$f(0.5) = \int_0^\infty \sqrt{t} e^{-t} \, dt$$

.

This function $f(x)$ is by no means the only possible extension of the factorial concept. You can construct infinitely many different continuous, infinitely-differentiable functions $f(x)$ that have the properties that $f(x) = xf(x-1)$ for all x and $f(x) = x!$ when x is a non-negative integer.

However, these other functions involve much more complicated definitions rather than a simple, elegant formula. The definition given above is the most "natural" (but not the only) way to extend the meaning of "factorial" to non-integer values.

Two questions now arise: why does the above definition have the property that $f(x) = xf(x-1)$, and secondly, the heart of your question, why does $f(0.5) = \sqrt{\pi/4}$?

The first question is answered through the technique of *integration by parts*. (I hope you've seen some calculus or this explanation won't mean much to you). If u and v are two functions, the product rule tells us that $(uv)' = u'v + uv'$, so $uv' = (uv)' - u'v$. Therefore, the integral of uv' is the same as the integral of $(uv)'$ minus the integral of $u'v$. The first integral is, by the fundamental theorem of calculus that says the integral of a derivative is the original function evaluated at the endpoints, uv evaluated at the endpoints of integration.

Applying this to $f(x)$ which is the integral of $t^x e^{-t}$, and letting $u = t^x$, $v' = e^{-t}$, we have $u' = xt^{x-1}$ and $v = -e^{-t}$. We therefore have

$$\begin{aligned} f(x) &= t^x(-e^{-t})\big|_0^\infty - \int_0^\infty xt^{x-1}(-e^{-t})dt \\ &= -\lim_{t\to\infty} t^x e^{-t} + 0^x e^{-0} + xf(x-1) \\ &= 0 + 0 + xf(x-1) \end{aligned}$$

(the first term can be shown to be zero through l'Hôpital's rule; intutitively, e^{-t} goes to zero much faster than t^x blows up).

Now, as to why a π is involved in the value of $f(0.5)$. We calculate the integral by making the substitution $x = \sqrt{t}$, so $t = x^2$ and $dt = 2xdx$:

$$\begin{aligned} f(0.5) &= \int_0^\infty \sqrt{t}e^{-t}dt = \int_0^\infty xe^{-x^2}2x\,dx \\ &= 2\int_0^\infty x^2 e^{-x^2}dx = \int_{-\infty}^\infty x^2 e^{-x^2}dx. \end{aligned}$$

(The last equality is by symmetry).

Next, we do an integration by parts with $u = x$ and $v' = xe^{-x^2}$ to transform this integral into

$$f(0.5) = \frac{1}{2}\int_{-\infty}^\infty e^{-x^2}\,dx.$$

The trick here is to square the integral and use y instead of x as the variable of integration in the second factor:

$$f(0.5)^2 = \frac{1}{4}\int_{-\infty}^\infty e^{-x^2}dx \int_{-\infty}^\infty e^{-y^2}dy.$$

This is the same as one quarter times the "double integral" of the function $e^{-x^2}e^{-y^2}$ over the entire xy-plane. Now, because of the properties of exponentials, this equals $e^{-(x^2+y^2)} = e^{-r^2}$ where r is distance to the origin. This function is *radially symmetric*; it depends only on distance to the origin and not on the angle. Therefore, on a small ring of radius r and small thickness Δr this function is roughly constant, with the value e^{-r^2}. The double integral over the ring is therefore approximately e^{-r^2} times the area of the ring, which is approximately $2\pi r\Delta r$: circumference times thickness.

If we add this up over a series of concentric rings, and take the limit as the thickness of each ring goes to zero, the errors in the approximations go to zero and we end up with

$$f(0.5)^2 = \frac{1}{4}\int_0^\infty 2\pi r e^{-r^2}dr = \pi/4$$

from which one sees that $f(0.5) = \sqrt{\pi/4}$.

One reason for the appearance of the π is that the properties of exponentials give you a radially symmetric function when doing the above calculation, and the π appears as part of the formula for the circumference of the rings on which that function is roughly constant.

Another reason the appearance of the π should not be surprising is that π is intimately connected to the properties of exponentials: when the exponential function is extended to complex numbers, its period is $2\pi i$ (in the sense that, if you add $2\pi i$ to the input, the output remains unchanged). This in turn is because $e^{2\pi i} = 1$, which is described in the answer to another question.

Solution to the Transcendental Equation $2^x + 3^x = 5$

Asked by B. Ryan, Brebeuf College on Wednesday Jan 10, 1996:

> Either I have forgotten or have not come across a non Newtonian (numerical approximation) to solve the following problem.
>
> $$2^x + 3^x = 5.$$
>
> Clearly the solution to the problem is that $x = 1$.
>
> By a Newton's method calculation involving first derivative, the soution is also easily obtained.
>
> What exact method, not involving an approximation is there to solve this problem.
>
> If available, solve the problem $4^x + 5^x = 100$
>
> solution: $x = 2.5843539862708$
>
> B. Ryan

Unfortunately, the solution to almost every transcendental equation (equation involving functions other then simple polynomials) can not be expressed as a combination of elementary functions, even if the equation itself can be.

Thus, while the equation $2^x + 3^x = 5$ happens to have a nice integer answer "1", there is no general formula for expressing the solution to the equation $a^x + b^x = c$ as a combination of elementary functions of a, b, and c.

There are certain special classes of equations for which the solution *can* be expressed as elementary functions. For an obvious example, the solution to the equation $a^x = b$ can be expressed as $x = \log(b)/\log(a)$. However, such special classes of equations are the exception rather than the rule; it is provable that no expression exists in the general case.

Followup question by Jeyprakash Michaelraj Fernando, India on November 5, 1996:

> I am wondering how to solve the equations of the type $a = x\sinh(b/x)$ (solve for x).
>
> Even by Iteration method, is there anyway by which I can make a guess for the initial value?
>
> THANKS in ADVANCE.

For a relatively uncomplicated equation like this, a binary search is often a good way to go.

First, one should figure out the general behaviour of the function to get an idea of where (and if!) a solution would be.

If we let $f(x) = x \sinh(b/x)$, it is easy to see that f is a continuous function everywhere except at $x = 0$. The limiting behaviour at $\pm\infty$ and at 0 can be found by calculating

$$\begin{aligned} \lim_{x\to 0} f(x) &= \infty \\ \lim_{x\to\pm\infty} f(x) &= b \end{aligned}$$

So, if your value of a satisfies $a > b$, you can solve the equation as follows:

1. Choose $x = A$ large enough that $f(x)$ is close to b and hence less than a. (You could either make a sophisticated estimate of how large to take x, or else just start with something like $x = 1$ and keep doubling it until it's large enough).
2. Choose $x = B$ small enough (close enough to 0) that $f(x)$ is close to infinity and hence greater than a. (Again, you could just start with $x = 1$ and keep halving it until it's small enough).
3. Now the intermediate value theorem guarantees that there's a solution between A and B. Let x be the halfway point $(A + B)/2$. Check if $f(x)$ is greater than or less than a. If it's greater than, you now know there's a solution between x and B. If it's less than, you now know there's a solution between A and x.
4. Either way, you've cut the size of the inteval in half. You can iterate this procedure, cutting the interval in half each time, until you have trapped the root within an interval of your desired accuracy.

This method is slow and inefficient compared to other methods, but is often a good one if you are having trouble finding an initial point.

The iteration method (finding a solution to an equation of the form $x = g(x)$ by forming the sequence $x, g(x), g(g(x)), \ldots$) will converge to a solution y if you start with x close enough to y *and* if $|g'(y)| < 1$. However, if $|g'(y)| > 1$, the iteration method will not converge to the solution y. This may be the problem you are having.

A better method is Newton's method. As long as $g'(y)$ is nonzero, and you start with x close enough to y, Newton's method will produce a sequence of numbers that converge to y, and will do so more rapidly than the bisection method. But if you're having trouble choosing an initial value that works, the bisection method can help you get close enough to the root, then you could switch to Newton's method to do the rest of the calculation more rapidly.

Finally, note that the above analysis of the function only shows that a solution exists when $a > b$. In the case $a \le b$ there are *no* solutions. You can prove this by showing that f is an increasing function when $x < 0$ and a decreasing function when $x > 0$. From this it follows that $f(x)$ can never get any smaller than the limiting value of b, so there are no solutions to the equation $f(x) = a$ if $a < b$.

In summary: the most important thing in problems of this type is to analyze the function's behaviour, using the tools of calculus. Figure out its limits as x approaches plus or minus infinity and as x approaches any points of discontinuity. Find where the function is increasing or decreasing. Using this, you can determine what part of the real number line a solution will lie on (if there is a solution). The bisection method can always be used to find the solution in this case; and you can use the faster Newton's Method once you're sufficiently close to the solution.

I should point out that most problems in numerical analysis are not as easy as this. For instance, you may have a function that dips below the x-axis only briefly, and is positive near both ends. Then you cannot use the simple analysis I described above to find solutions to $f(x) = 0$, and unless you're fortunate enough to know at least one x value where $f(x)$ is negative and at least one where $f(x)$ is positive, you cannot use the bisection method.

In these cases you have to start with Newton's Method or something more sophisticated, and entire books can be (and have been) written on how to choose appropriate initial points, what effect small errors in calculations have on the final answer, etc. Any textbook with a title like "Numerical Analysis" should be able to explain these issues in much more depth.

Limit of the Sequence $a(n) = \cos(a(n-1))$

Asked by John Koehler, student, High Technology High School on February 17, 1997:

> I have a math question that you may or may not be able to answer. If you can't answer it please refer me to someone who you believe can. I am in calc2 so your answer can be somewhat complex.
>
> My question is:
>
> I have fallen upon the sequence $a(n) = \cos(a(n-1))$.
>
> My question is, what does this sequence converge on? I know that it goes towards a number that is about $0.73\ldots$. In fact I can find the value to about 12 decimal places using a calculator but I am interested in finding out what this number is (in terms of other things like integers, π, e, or other things).
>
> I understand if you don't have time to answer but please reply.
>
> Thank you for your time,
>
> John Koehler

The number this sequence converges to is the unique solution to the equation $x = \cos x$. In other words, it is the place where the graph of the function $f(x) = x - \cos x$ crosses the x-axis. Let's call this number z.

First, let me explain how we know such a number z exists, and why it is unique.

Its existence follows from the intermediate value theorem: $f(0) = 0 - 1 < 0$ while $f(\pi/2) = \pi/2 - 0 = \pi/2 > 0$, and f is a continuous function, so there must be some number z in between 0 and $\pi/2$ with the property that $f(z) = 0$.

The reason this number z is unique is that the function f is non-decreasing: its derivative is $f'(x) = 1 + \sin(x) \geq 1 - 1 = 0$. In order for it to have two roots, it would have to either be constant between them (which it clearly is not; there is no interval on which $x - \cos(x)$ stays constant), or else it would have to be decreasing somewhere (if it increases after passing through the first root, it would have to decrease again to get back to zero), and this can't happen because f is non-decreasing.

This explains how the number z is uniquely defined by the equation $\cos(z) = z$.

The reason the sequence converges to this number z is as follows. Let's look at how close $a(n)$ is to this number z: the absolute value of the difference is $|a(n) - z| = |\cos(a(n-1)) - z|$. Since $z = \cos(z)$, this is the same as $|\cos(a(n-1)) - \cos(z)|$.

By the mean value theorem, any expression of the form $g(b) - g(a)$ equals $g'(t)(b - a)$ for some number t in between a and b (assuming g is differentiable on the interval $[a, b]$). I hope you have seen this theorem in your calculus class; if not, let us know and we can explain further.

Therefore, $|a(n) - z| = |\cos(a(n-1)) - \cos(z)| = |\sin(t)||a(n-1) - z|$.

Now, regardless of what the first term of the sequence is, the second and all subsequent terms are the cosine of something and hence are between -1 and 1. So too is the number z, so the number t above will also be between -1 and 1. On this interval the sine function is strictly increasing, so $|\sin(t)| < |\sin(1)|$.

Applying this inequality to the above expression, we get

$$\begin{aligned} |a(n) - z| &< |\sin(1)|\,|a(n-1) - z| \\ &< |\sin(1)|^2\,|a(n-2) - z| \\ &< \ldots \\ &< |\sin(1)|^n\,|a(0) - z| \end{aligned}$$

Since $|\sin(1)| < 1$, this goes to zero as $n \to \infty$, proving that the sequence converges to z.

This is all a special case of a general theorem called the *contraction mapping principle*. Any time you have a function F. an interval $[a, b]$, and a number $e < 1$, with the property that (a) $F(x)$ is in the interval whenever x is, and (b) $|F(x) - F(y)| < e|x - y|$ for all x and y in the interval, then there is guaranteed to be a unique number z (called a *fixed point*) in $[a, b]$ for which $F(z) = z$, and, if you start with any number x in $[a, b]$ and look at the sequence $x, F(x), F(F(x)), \ldots$, it converges to z.

Now, finally, for the main point of your question: is there any simpler way to express the number z (as a combination of familiar numbers, like integers, e, and π), other than simply calling it "the solution to the equation $\cos(z) = z$"?

The answer is no. This number z cannot be expressed any more simply, in terms of "familiar" numbers. This is true for the roots of almost any transcendental equation; for more information, see the answers to other questions on this topic.

If you were in a context where you needed to work with this number a lot, you'd simply give it a name, in much the same way as people gave names to the numbers π and e.

Solving a Quadratic with Non-Constant Coefficients

Asked by Alex Pintilie, teacher, Bayview Glen School on April 25, 1997:

> I put in a test the following question.
>
> > Prove that the following equation has no roots: $\sin x = x^2 - 4x + 6$.
>
> (I had in mind a graphical solution.)
>
> A student gave me the following solution:
>
> > $x^2 - 4x + (6 - \sin x) = 0$. $\Delta = 4 - (6 - \sin x) = \sin x - 2$ has to be ≥ 0 to have solutions. $\sin x$ cannot be > 1, so the equation has no solutions.
>
> He is "sort of right" but I am "afraid" of allowing students to solve for x with respect to x. What do you think?

Your student's answer is perfectly correct, except that there's a step missing: the discriminant of the quadratic is $b^2 - 4ac = 4^2 - 4(1)(6 - \sin x)$, which is $4(\sin x - 2)$ rather than just $\sin x - 2$. However, the extra factor of 4 does not affect the sign of the discriminant.

The answer is correct because, if x, a, b, and c are any four real numbers which satisfy the relationship $ax^2 + bx + c = 0$ (*), then the discriminant $b^2 - 4ac$ must be non-negative. If there were a solution x to your original equation, then the four numbers x, $a = 1$, $b = -4$, and $c = 6 - \sin x$ would satisfy the quadratic relationship (*), so the discriminant would have to be non-negative. However, the student has shown that in fact the discriminant is negative.

Here's another way to think about the student's answer. Suppose your original equation had a solution $x = r$. That would mean $r^2 - 4r + (6 - \sin r) = 0$. Now, think of the quadratic $x^2 - 4x + (6 - \sin r) = 0$ where r is just a fixed constant and x is a variable not necessarily related to r. Now you have a true quadratic function of x being set equal to zero. The student has shown that this quadratic has no real roots (no matter what value the constant r has, the discriminant will always be negative). Since it has no roots at all, that proves that in particular $x = r$ is not a root, and hence it cannot be the case that $r^2 - 4r + (6 - \sin r) = 0$, no matter what r is.

As for the question of "solving for x in terms of x": if you carry out the derivation of the quadratic formula, it shows that if x, a, b, and c are any four numbers related by the equation $ax^2 + bx + c = 0$, then they are also related by the equation

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

This is just a re-writing of the original equation by a process of completing the square.

Now, if the numbers a, b, and c happen to depend on x, as they do in your example, it is still legitimate to rewrite the equation as

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

And, in the case of your example (where $a = 1$, $b = -4$, and $6 - \sin x$), one can note that the term inside the square root is always negative, so the rewritten equation has no real solutions, so the original equation has no real solutions either.

The only difference between this case and the case where a, b, and c are constants is that here the rewritten form of the equation is not a *solution* for x, merely a different equation involving x. It is therefore of no use, except in theoretical arguments like this. However, in the case where a, b, and c are constants, the rewritten form of the equation is actually a solution for x in terms of a, b, and c.

Therefore, although the rewritten form of the equation is always valid, it is only a solution for x in the case in which a, b, and c are constants independent of x. It is usually only in this case that it is useful. Typically, students trying to apply it in other cases will do so incorrectly, but the student who answered your test question is a welcome exception!

Followup Question by Alex Pintilie, teacher, Bayview Glen School on April 29, 1997:

> This is just a follow up on a question I asked earlier about a "non-standard" solution to a quadratic equation problem. Thank you for your answer. I gave my student a printout of your argument and full marks for his solutions. However, just for the sake of playing "devil's advocate" here's another simmilar hypothetical problem and solution.
>
> *Define $f(x) = x^2 - 4x + 6$ on the domain $\{x : x > 3\}$. Show that $f(x) = 0$ has real solutions.*
>
> Rewrite $f(x) = x^2 - 2x + 6 - 2x$. $\Delta = 4 - 4(6 - 2x) = 8x - 20 > 0$ since $x > 3$. Since $\Delta > 0$, the quadratic has solutions.
>
> Since there is nothing wrong with $c = 6 - \sin x$, there should be nothing wrong with $c = 6 - 2x$.
>
> Thank you again for answering my question.

The flaw in this second argument is the following. The fact that $\Delta > 0$ means that the quadratic has solutions *in the sense that there is a number y for which $ay^2+by+c = 0$, i.e., $y^2-2y+6-2x = 0$.* However, there is no reason why this root y should necessarily be the same thing as the number x that occurs in the c coefficient, so it is not legitimate to conclude from this that there is a solution x to the equation $x^2 - 2x + 6 - 2x = 0$.

However, the first argument was correct because it showed that there was no solution y at all to the equation $y^2 - 4y + 6 - \sin x = 0$, and since there is no solution at all, in particular $y = x$ cannot be a solution.

Perhaps it might be useful to rephrase this in terms of the quadratic formula. The original quadratic equation $ax^2 + bx + c = 0$ can be rewritten as

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

This rewriting is always valid.

Now, if a, b, and c are constants, this gives a *solution* for x in terms of a, b, and c (or rather, two solutions); these solutions are real numbers if and only if $b^2 - 4ac \geq 0$. Therefore, in the case when a, b, and c are constants, the quadratic has real solutions if and only if $b^2 - 4ac \geq 0$.

However, if a, b, or c is a function of x, then the rewritten equation is simply a new, more complicated equation involving x. If $b^2 - 4ac < 0$ then clearly the equation can have no real solutions because negative numbers do not have real-valued square roots. The reverse implication, though, does not hold: if $b^2 - 4ac \geq 0$, that does not necessarily mean there *is* a solution, only that a solution can no longer be ruled out on such basic grounds.

To put it yet another way: if $b^2 - 4ac \geq 0$, then you can certainly define a real number y by the formula

$$y = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

but there's no guarantee that this number y can be made to be the same number as the number x that occurs in the coefficients.

When the coefficients are constants there is no such requirement; so in the constant coefficient case you always get solutions when $b^2 - 4ac \geq 0$.

In the non-constant-coefficient case, you know there cannot be solutions if $b^2 - 4ac < 0$, but in the case $b^2 - 4ac \geq 0$ there is nothing about the rewritten equation that allows you to easily tell if it has solutions or not.

I hope this helps clear things up and I hope this makes for some useful discussions with your students.

Solution to a Functional Equation

Asked by Aryan Ghirati, student, Kherad on October 6, 1997:

> Is there a function f from $\mathbf{N}$ (the set of natural numbers) to itself, such that such that f^2 (the composition of f with f) sends every n to n^n?

Yes, there is such a function. All such functions arise through the following type of construction:

Let C denote the set $\{1^1, 2^2, 3^3, \ldots\}$. Subdivide the rest of the natural numbers (those that are not in C) into two disjoint sets A and B, in any way you like, except that you need to have infinitely many numbers in each of the two sets. (For example, you could list the numbers that are not in C: 2,3,5,6,7,.... Then let the set A consist of the first number in this list, the third number in this list, and so on, while B consists of the second number in this list, the fourth number in this list, and so on).

Enumerate the numbers in A, calling them $a_1, a_2, \ldots$. Similarly, enumerate the numbers in B, calling them $b_1, b_2, \ldots$.

Now define f on the sets A and B by defining $f(a_i) = b_i$ and $f(b_i) = {a_i}^{a_i}$. (Note that f maps A to B and maps B to C).

Define f on the set C by setting $f(1) = 1$ and, for all $n > 1$, define $f(n^n) = f(n)^{f(n)}$.

If n is in A or B, this definition is well-defined, for we have already defined $f(n)$, and the above formula tells us how to define $f(n^n)$.

If n is in C, the way we make this definition correct is to think of it as an inductive definition: assume we have already defined $f(m^m)$ for all $m < n$ and are now trying to define $f(n^n)$. In the case in which n happens to be in C, that means $n = m^m$ and it is easy to verify that $m < n$, hence $f(m^m)$ has already been defined, and we get an inductive definition of $f(n^n)$ in terms of it.

The function f has now been defined for all natural numbers. It remains to show that $f(f(n)) = n^n$ for all n. There are three cases.

Case 1: n is in A. Then $n = a_i$ for some i. We then have $f(f(n)) = f(b_i) = a_i^{a_i} = n^n$ as required.

Case 2: n is in B. Then $n = b_i$ for some i. We then have $f(f(n)) = f(a_i^{a_i}) = f(a_i)^{f(a_i)} = b_i^{b_i} = n^n$ as required.

Case 3: To prove that $f(f(n)) = n^n$ for all n in C, suppose to the contrary that there are some n in C for which $f(f(n)) \neq n^n$. Let $n = N$ be the smallest of these. We know $N > 1$, since $f(f(1)) = 1 = 1^1$. We also know N is in C, so $N = m^m$ for some m, and clearly $m < N$ (the only time $m = m^m$ is when $m = 1$, and we know that our number N is greater then 1).

Either m is in A, or it is in B, or it is in C but is smaller than N, and in all three of these cases we already know $f(f(m)) = m^m = N$. It follows that $f(f(N)) = f(f(m^m)) = f(f(m)^{f(m)}) = f(f(m))^{f(f(m))} = N^N$. This contradicts our choice of N as being the smallest counterexample to $f(f(n)) = n^n$, showing that no counterexample exists, and hence f has the desired property.

You can also prove that *all* functions with this property are given by a construction of the above form. Assume we have such a function f. The first thing to note is that f must satisfy the property $f(n^n) = f(n)^{f(n)}$ for all n, since $f(n^n) = f(f^2(n)) = f^2(f(n)) = f(n)^{f(n)}$.

Now, let C be the same set as above, and divide the numbers that are not in C into two classes, A and B, where A is the set of x not in C for which $f(x)$ is not in C, and B is the set of x not in C for which $f(x)$ is in C.

One can prove that f sets up a one-one correspondence from A to B, as follows: if x is in A, we know $f(x)$ is not in C, but $f(f(x)) = x^x$ which is in C, proving that $f(x)$ is in B. This proves that f maps A into B.

This mapping is one-to-one: if $f(x) = f(y)$, then $f(f(x)) = f(f(y))$, so $x^x = y^y$, and it's easy to verify that for natural numbers this only happens when $x = y$.

This mapping is onto: every element $b \in B$ is obtained as $f(a)$ for some $a \in A$. This is because $f(b)$ is in C, so $f(b) = a^a$ for some a, and $b^b = f(f(b)) = f(a^a) = f(a)^{f(a)}$ from which it follows that $b = f(a)$.

So, if you label the elements of A as $a_1, a_2, \ldots$ and define $b_i = f(a_i)$, then f is precisely the function defined above, since $f(a_i) = b_i$ (by definition), $f(b_i) = f(f(a_i)) = a_i^{a_i}$ (by the argument of the previous paragraph), and for every element $n^n \in C$, $f(n^n) = f(n)^{f(n)}$. This is precisely the definition given above.

Largest Possible Number?

Asked by Carlos Gomez on January 24, 1998:

> What is the name of the Highest Possible Number? Is it called a "Gugle"? I know there's a specific name, and I don't mean "Infinity".
>
> Thank You, Carlos A. Gomez

There is no such thing as a highest possible number. No matter what number you have, there is always a larger one. (For example, you could always add 1 to your number to get a larger number).

What you are intending to ask is "what is the largest number that anyone has ever decided to give a specific name to?" It is very important to understand that this is a *completely different question* from the one you asked, because it is a question about human culture not about mathematics.

The largest number that has a commonly-known specific name is a "googleplex", which is a 1 followed by a googol zeros, where a "googol" is 10^{100} (a 1 followed by 100 zeros).

However, there would be nothing stopping you from giving a special name to a still larger number (such as a googleplex plus 1), and then *that* would become the largest named number once the term became commonly known.

In summary, then: the mathematical question "what is the highest possible number" has no answer, because there is no such thing. But the sociological question "what is the largest number that anyone has ever decided to give a specific name to, a name which has become commonly known" is, for now, a "googleplex" (until someone decides to coin a phrase for a still larger number and it catches on and becomes commonly known).

Calculating Digits of Pi in Other Bases

Asked by an anonymous poster on December 12, 1997:

> Could you please give us a formula for figuring out what pi is in base 6? We have the numbers 3,0,5,0,3,3,0,0,5,1,4,5, etc. but we do not know exactly how to obtain these numbers.

Digits of pi are generally calculated by representing pi as the sum of an infinite series, then evaluating that sum to as many terms as are necessary to get the desired number of digits.

One way is to use the formula $\mathrm{Arctan}(x) = x - x^3 + x^5/5 - x^7/7 + \ldots$, which is valid for $x \leq 1$. Plugging in $x = 1$ is not much use because, although you get a formula $\pi/4 = 1 - 1/3 + 1/5 - 1/7 + \ldots$, the series converges *very* slowly and you have to take thousands of terms to get even a few digits of accuracy.

A faster way is to use a formula such as $\pi = 16\mathrm{Arctan}(1/5) - 4\mathrm{Arctan}(1/239)$. You evaluate each of these arctangents using the above series, taking enough terms of the series to get you to within the accuracy you desire. In your case, you would do the arithmetic in base 6 to get digits of π in base 6.

Another, more recent, series is

$$\pi = \sum_{k=0}^{\infty} \frac{1}{16^k} \left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6} \right)$$

which has the advantage that in base 16 there turns out to be a clever way to use this formula to calculate far-out digits of π without having to do high-precision arithmetic and without having to compute all the earlier digits first. That's only a help in base 16 and a few other bases, though.

For example, if you were to use this latter series to calculate the first five base-6 digits of π, you would say that you want π to within an accuracy of 6^{-6} (which is about 0.00002). Note that the k-th term in the above series is smaller than $1/16^k$, so if you add up the terms from $k = n+1$ onwards the most you will get is $1/16^{n+1} + 1/16^{n+2} + \ldots$ which, by the formula for summing a geometric series, is $\frac{1}{15\times 16^n}$. This, then, is the maximum error you could possibly introduce by taking only the first n terms of the series.

If you take $n = 3$, the above estimate gives you an error of at most 0.0000163 which is smaller than your desired accuracy. Therefore, all you need to do is take the first three terms of the series. You would evaluate the terms when $k = 0$, when $k = 1$, when $k = 2$, and when $k = 3$ and add them up, doing the calculations in base 6.

Tetrahedral and 4-Tetrahedral Numbers

Asked by Warren Kendrick on November 4, 1997:

> The tetrahedronal numbers: 1, 4, 10, 20, 35, ... are derived by adding the triangular numbers: $1 + 3 + 6 + 10 + 15 + \cdots + (n)(n+1)/2$. At one time I derived a formula for the sum through the mth term of the tetrahedronal numbers, and also the sum of the numbers for a 4-tetrahedron: 1, 5, 15, 35, 70, ..., which are of course the sum of the tetrahedronal numbers. (Tetrahedronal numbers are best described by placing larger and larger triangular layers of 3-spheres under the previous triangular layer of spheres, forming a tetrahedronal pyramid.) Can you provide these formulas? Thanks.

These numbers are all expressible as binomial coefficients. If we let $T_2(n)$ denote the nth triangular number, $T_3(n)$ the nth tetrahedral number, and so on, then

$$T_k(n) = \binom{n+k-1}{k} = \frac{(n+k-1)!}{(n-1)!k!}.$$

You can prove this by induction noting that each $T_k(1)$ equals 1, that the formula is correct when $k = 2$, and the difference $T_k(n+1) - T_k(n) = T_{k-1}(n+1)$.

You could also prove this by a more "brute-force" approach, using the formulas for the sums of powers:

$S_1(n) = 1 + 2 + 3 + \cdots + n = n(n+1)/2$

$S_2(n) = 1^2 + 2^2 + 3^2 + \cdots + n^2 = n(n+1)(2n+1)/6$

$S_3(n) = 1^3 + 2^3 + 3^3 + \cdots + n^3 = n^2(n+1)^2/4$

and so on. To arrive at these formulas you can again proceed by brute force (make the intelligent guess that $S_k(n)$ should be a polynomial of degree $n+1$ and plug in enough values to solve for the coefficients), or else employ a trick like the following:

Write $(m+1)^{k+1} = m^{k+1} + (k+1)m^k$ plus terms involving lower powers of m. When you add up the LHS as m ranges from 0 to n, you get $S_{k+1}(n+1) = S_{k+1}(n) + (n+1)^{k+1}$. When you add up the RHS, you get $S_{k+1}(n) + (k+1)S_k(n)$ plus terms involving $S_{k-1}(n)$ down through $S_1(n)$. This gives you

$$S_{k+1}(n) + (n+1)^{k+1} = S_{k+1}(n) + (k+1)S_k(n) + \text{lower order terms}$$

The terms $S_{k+1}(n)$ cancel, and you are left with a formula that expresses $S_k(n)$ in terms of $S_{k-1}(n)$ down through $S_1(n)$.

For example, here is how to use the above technique to arrive at the formula for $S_3(n)$:

$$\begin{aligned}
(m+1)^4 &= m^4 + 4m^3 + 6m^2 + 4m + 1 \\
S_4(n) + (n+1)^4 &= S_4(n) + 4S_3(n) + 6S_2(n) + 4S_1(n) + 1(n+1) \\
4S_3(n) &= (n+1)^4 - 6S_2(n) - 4S_1(n) - n - 1 \\
S_3(n) &= (1/4)[(n+1)^4 - 6n(n+1)(2n+1)/6 - 4n(n+1)/2 - n - 1]
\end{aligned}$$

and simplifying gives the above formula. (Note that the reason $n+1$ appears at the end of the second line is that we are adding up the first line as m ranges from 0 to n; that means the final constant term "1" is added up $n+1$ times.)

Now that you have formulas for $S_k(n)$, it is easy to find formulas for the tetrahedral numbers. For example,

$$\begin{aligned}
T_2(n) &= \sum_{m\le n} m(m+1)/2 \\
&= (1/2)\sum_{m\le n} m^2 + (1/2)\sum_{m\le n} m \\
&= (1/2)n(n+1)(2n+1)/6 + (1/2)n(n+1)/2 = n(n+1)(n+2)/6.
\end{aligned}$$

Existence of Shapes with Irrational Dimensions

Asked by Jon Cypryk, student, A.N. Myer on October 27, 1997:

> We would like to pose a question which we have been unable to reach an agreement on. The question is "can a six-sided, 3-D shape exist with a combined surface area of 600 square units? Two opposite sides must have an area of 100 each, two opposite sides with an area of 150 each, and two opposite sides with an area of 50 each."
>
> The tentative answer given is a shape with measurements of a=square root of 75, b=150 divided by the square root of 75, and c=50 divided by the square root of 75. Then $a \times b = 150$ square units, $b \times c = 100$ square units, and $a \times c = 50$ square units.
>
> We understand that this formula works out mathematically because the square root of 75 is rounded off for all intents and purposes because it would be insane not to since it can bring you infinitely close as possible to an area of 600 square units.
>
> However, due to an intense desire to prove or disprove the above statement, we are being precise, exact and very stubborn to the point.
>
> Is it correct that if the length of one of the sides cannot exist (eg., a square of 75), then we could not find an area of one sides equaling exactly and precisely 100 square units? And if this is so then can the shape described in the opening question exist with an exact area of 600 square units?

> The item in question would seem to be the length of the a side equaling square of 75.
>
> The two opposing beliefs are:
>
> - the belief that there is no true or exact number for the square root of 75 since it is irrational and continues on forever with no repeating pattern. Therefore side a cannot exist if we choose to be stubborn and exact to the point, which we must in order to get the exact result of 600 without rounding off.
>
> - the belief that even though we cannot get to or see the end of the square root of 75, it it plausible that it can still exist and therefore it is plausible that side a can exist with an exact measurement of square root of 75.
>
> An answer to this would be greatly appreciated if possible to help us prove or disprove the original question.
>
> Thank-you in advance for your help.
>
> Jon Cipryk

The real question you are asking is "do irrational numbers exist?" They most certainly do. And so yes, the shape described above does exist.

Several issues need to be addressed to clear up the confusion. First, there are different kinds of numbers. One kind of number is concept of "natural number": the sort of number used to measure "how many". If you were to ask the question "does there exist a number between 1 and 2?", the answer would be "no" if you were referring to the kind of numbers used in counting. For example, it is not possible to press a computer key more than once but less than twice.

However, that does not mean that the number 3/2 does not exist! It just means that it isn't the sort of number used in counting. It exists as a "rational number": a ratio of two integers.

In the same way, a number like the square root of 75 does not exist in the context of rational numbers (just as "half of three" does not exist in the context of the integers). But it does exist in the context of a different number system called the "real numbers" (just as "half of three" does exist in the context of rational numbers).

There are several ways to rigorously define real numbers. One way is to define a real number to be a sequence of rational numbers. So, for example, the sequence of rational numbers 8, 8.6, 8.66, 8.660, 8.6602, ... defines the real number $\sqrt{75}$. Note that no individual number in that sequence defines $\sqrt{75}$ (which is what you were getting at when you said that no finite decimal exactly equals $\sqrt{75}$); however, the *entire sequence* taken together defines $\sqrt{75}$.

Another way to define real numbers is to define a real number as a partitioning of the rational numbers into two sets, where everything in the first set is less than everything in the second set. (Intuitively, such a partition corresponds to a location on the number line: the place where the first set ends and the second set begins). Now, $\sqrt{75}$ corresponds to a perfectly well-defined partition of the rational numbers: for each positive rational number r, either $r^2 < 75$ or $r^2 > 75$, and this distinction lets us separate the rationals into two classes, thereby defining a real number if you interpret "real number" as meaning "partition of the rationals into two sets with the appropriate properties".

Each of these definitions is quite abstract. (If you completely understand the previous two paragraphs, you should consider yourself exceptionally gifted in mathematics and I'd encourage you to consider it as a career). Therefore, they are not usually taught until about the third year of an undergraduate program. However, the important thing is that there are such things as "real numbers", that can be rigorously defined (though the definition is abstract and difficult), and within this collection of real numbers there is one whose square is 75. Therefore, the square root of 75 exists.

It's important to realize that these "real numbers" are not just an artificial mathematical construction but are precisely the kind of number system relevant for length measurements (just as the natural numbers are the kind of number system relevant for counting). As a consequence, the square root of 75 exists not just as an abstract mathematical entity, but as a real geometrical length.

One good way to see that some real, physically existing lengths can only be measured by irrational numbers is to think of a square with side length 1. The diagonal of this square (a length that clearly "exists") has length $\sqrt{2}$, which is an irrational number.

One final confusion that arises is this: there's a temptation to forget the distinction between a *number* and a *decimal representation of a number*. The number 75, for example, is an abstract entity that exists in its own right quite independently of the fact that can be written as the sum $7 \times 10 + 5$ so that we can write it down as a 7 followed by a 5. When you come across a number like $\sqrt{75}$ and observe that it cannot be written down as a finite sum of the above form, meaning that there's no finite decimal representation for it, that doesn't mean the number itself fails to exist. It just means it's a number that happens not to have a finite decimal representation.

What are the Origins of Number Systems?

Asked by Casey Sloan, student, The Lovett School on March 17, 1997:

> Who came up with the idea of number systems and for what reasons?

The origins of numbers date back to the Egyptians and Babylonians, who had a complete system for arithmetic on the whole numbers $(1, 2, 3, 4, \ldots)$ and the positive rational numbers.

The Greeks at the time of Pythagoras knew that these number systems (whole numbers and ratios of whole numbers) could not completely describe everything they wanted numbers to describe. They discovered that no rational number could describe the length of the diagonal of a square whose sides were of length 1. They called such lengths "irrational", recognizing that some other kind of number system would be needed in order to describe them, but not knowing what it would be. They did not pursue the matter, for they viewed whole numbers with such awe that anything not expressible in terms of whole numbers was distrusted by them as contrary to nature.

These number systems evolved somewhat during the Middle ages with the notable addition by the Hindus of a convenient notation for zero and negative numbers, concepts which previously had been difficult to deal with due to the lack of notation. The properties of the "real number system" (consisting of both rational and irrational numbers) began to be understood in the 1600's with the development of calculus, and by the end of the 1800's mathematicians such as Dedekind and Cantor were giving rigorous mathematical definitions of this number system, putting it on equal footing with the whole numbers and rational numbers.

It wasn't until the early 1800's, however, that the abstract structure of these number systems was studied. This new area of math, like many other areas of math, arose from a creative new way to answer an old question: how to find the roots of a polynomial (those numbers which, when substituted into it, give zero).

Much was known about polynomials of degree (highest power) less than 5. Italian mathematicians had solved for the roots of the 3rd and 4th degree polynomials in the 1500's. These solutions were always expressible in terms of "radicals" or nth roots of numbers. For a long time no one knew how to solve a general 5th degree polynomial for its root.

Polynomials of lower degrees were still of interest though. In search of a deeper understanding of them, Gauss studied quadratic (2nd degree) polynomials. Through his work, he found that the objects he was considering were related to each other in much the same way that numbers are related under addition or multiplication. In modern terms, he was considering "finite group structures": finite sets which are essentially like a number system, but with only one operation. In many of the groups which he worked with the order in which the operation was performed doesn't matter: $a \cdot b = b \cdot a$. Groups in which the operation commutes in this way are called *abelian groups*. It is believed that Gauss may have been one of the first to have a rough understanding of the structure of finite abelian groups.

Also related to the study of polynomials is the "theory of substitutions" studied by Lagrange, Vandermonde, and Gauss. A substitution is where the variable of the polynomial is replaced with a different expression (such as a new variable plus a constant). It is possible sometimes to make the "right" substitution and turn a very complicated polynomial into something much easier to handle. This led to the study of the permutations of a set. Also studied by Ruffini and Cauchy, the permutations of a set form a group structure as well, though in this case the order of operation matters and therefore the groups are non-abelian.

Any discussion of the study of polynomials and number systems incomplete without a mention of Galois. He was the first to fully understand the connections between these finite number systems and the behavior of the roots of polynomials. It follows from his work that there is no "nice" formula for the roots of some 5th degree polynomials. While he died in his early 20s in a duel, his work (which was allegedly written in a letter and sent to a friend the day before he died) is still one of the cornerstones of the study of number systems.

The Hypercomplex Numbers

Asked by Marcello Praca Gomes da Silva, teacher, U.E.S. on July 7, 1997:

> Hello. I would like to know about hypercomplex numbers (of the form $a + xi + yj$), their properties and rules (addition, subtraction, multiplication, division, etc). In what areas are they useful ? Give me some examples if possible.

The hypercomplex numbers are a generalization of the complex numbers. They were created in an attempt to describe certain geometric operations in spaces with a dimension higher than the 2-dimensional plane.

Operations on the complex numbers can be used to describe many of the geometric operations on the plane. For instance multiplication by a real number corresponds to a scaling of the plane. Multiplication by complex numbers with a modulus ("length") of 1 corresponds to a rotation of the plane. Adding complex numbers corresponds to translation of the plane. More examples are available on the *geometry and imaginary numbers* page.

It was wondered for some time whether this could be generalized to 3-dimensional space (that is, to numbers of the form $a + xi + yj$, which is what you were asking about). It is known now that this is not possible. The only dimensions in which there are hypercomplex numbers which allow for a notion of division are dimensions 4 and 8. These hypercomplex numbers are called the *Quaternians* and *Octonians* respectively (the Quaternians are sometimes called *Hamilton Numbers* and the Octonians are often called *Cayley Numbers*).

In both cases these number systems are unique – the only numbers with these geometric properties in dimension 4 are the Quaternians and in dimension 8 are the Cayley Numbers. Also, with increases in dimension, we can not maintain some of the properties which make number systems simple. In both the Quaternians and the Cayley numbers, multiplication is non-commutative.

That is that if a and b are two numbers, then $(a)(b)$ need not be the same number as $(b)(a)$. In that case of the Cayley Numbers, multiplication is not even associative. By this we mean that if a, b, and c are three Cayley numbers, then it is not always true that $(ab)c$ is the same number as $a(bc)$ (here the parentheses tell us which operation to perform first).

While the definition of the Cayley Numbers is quite elaborate, the Quaternians can be defined without too much trouble. There are three Quaternians which, together with the Real numbers, generate all of the others. They are given the names i, j, and k. A typical example of a Quaternian is $1 + 3.3i + 9j - 1.5k$.

Addition is defined as you would define addition on a polynomial with variables i, j, and k. For example, the sum of $1 + 5j$ and $2 + 4j + 13k$ is $3 + 9j + 13k$.

Multiplication is also defined just like multiplication on polynomials except that we are allowed to make certain simplifications: $i^2 = j^2 = k^2 = ijk = -1$. We can derive other relations, such as $i = jk$, from the ones given above. Thus $(1+5j)(2+4j+13k)$ equals $(2+4j+13k)+(10j+20j^2+65jk) = -18+65i+10j+13k$.

The addition and multiplication of hypercomplex numbers are always defined in the same way, regardless of the dimension. What defines the behavior of the system is the rules for simplification. Other, more obscure hypercomplex number systems exist and are generally based around the multiplication of certain types of matrices. These, however, fail to have a notion of "division" and are more difficult to work with.

Why is the Product of Negative Numbers Positive?

Asked by an anonymous poster on March 18, 1997:

> I'm helping a 7th grader with things like: a plus times a plus equals a plus, a minus times a plus equals a minus, and a plus times a minus equals a minus. All OK. But when I tell him a minus times a minus equals a plus he says WHY? (sorry about yelling).
>
> I won't feel bad if you don't answer this. No textbook and nobody has the faintest idea. But just in case you do answer, please remember it's a 7th grader who wants to understand, not to mention yours truly.

The answer has to do with the fundamental properties of operations on numbers (the notions of "addition", "subtraction", "multiplication", and "division"). Your 7th grader's question is an important and fundamental one (which I am both surprised and sorry that he has not been able to find an answer for yet).

Each number has an "additive inverse" associated to it (a sort of "opposite" number), which when added to the original number gives zero. This is in fact the reason why the negative numbers were introduced: so that each positive number would have an additive inverse.

For example, the inverse of 3 is -3, and the inverse of -3 is 3.

Note that when you take the inverse of an inverse you get the same number back again: "$-(-3)$" means "the inverse of -3", which is 3 (because 3 is the number which, when added to -3, gives zero). To put it another way, if you change sign twice, you get back to the original sign.

Now, any time you change the sign of one of the factors in a product, you change the sign of the product:

($-$something) $\times$ (something else) is the inverse of (something) $\times$ (something else), because when you add them (and use the fact that multiplication needs to distribute over addition), you get zero.

For example, $(-3) \times (-4)$ is the inverse of $(3) \times (-4)$, because when you add them and use the distributive law, you get $(-3) \times (-4) + (3) \times (-4) = (-3+3) \times (-4) = 0 \times (-4) = 0$.

So $(-3) \times (-4)$ is the inverse of $(3) \times (-4)$, which is itself (by similar reasoning) the inverse of 3×4.

Therefore, $(-3) \times (-4)$ is the inverse of the inverse of 12; in other words, the inverse of -12; in other words, 12.

The fact that the product of two negatives is a positive is therefore related to the fact that the inverse of the inverse of a positive number is that positive number back again.

The answer to this question is accessible to a 7th grader (and should, in my opinion, be explained as part of every student's arithmetic classes). However, as an aside, he may be interested to know that more advanced versions of this question are studied at a university level: there is a subject called Abstract Algebra (usually only covered in a junior or senior level undergraduate university course) which studies the properties of operations on numbers in complete generality, even in contexts that have nothing to do with numbers at all. Even in such general, non-numerical contexts, the property that the product of two negative things is positive still holds.

Followup Comment by Buzz Breedlove on May 9, 1997:

> This is a comment on your answer to the question: "Why is a negative number times a negative number a positive number?" As a volunteer teacher for a pre-algebra class of sixth graders, I addressed the same question with the following practical demonstration. I randomly handed students each a bunch of red and black checkers. I announced that the blacks (hypothetically) represented each correct answer the student had given during the class. The red checkers represented (hypothetically) their wrong answers. I told them that for each black checker, I owed them a dollar, and for each red checker they owed me a dollar. They excitedly calculated their respective balances. I then advised the students that my accounting had been wrong and I had incorrectly given each student more red checkers than I should have given them. I went from student to student, taking back (subtracting) n red checkers. After each example, I had the student recalculate their balance. For each red checker (-\$1) subtracted, the students realized their balance increased by \$1. After just one example, all the students cheered in unison with the joy of understanding subtracting negative numbers. I then subtracted 2 red checkers three times from the next student. Again the students cheered realizing that subtracting 2 red checkers three times was like adding six to the balance sheet. They then understood that $-\$2 \times -3 = +\$6$. After this demonstration, students used negative numbers in their algebra with understanding.
>
> From my experience at Cosumnes River Elementary School Rancho Murieta, CA Ms. Lung's sixth-grade class of 1996
>
> Do you have a place to make comments like these?
>
> Buzz Breedlove breedlov@calweb.com

Thank you for your comments; I have placed them as followup comments to the question.

Followup question by Ms. White, Community College on October 3, 1997:

> About that question: why is a negative number times a negative number equal to a positive number? I think about addition when multiplying. That is $3x4 = 12$ and therefore $4+4+4 = 12$, therefore, $-3 \times -4 = 12$ because $-(-4) - (-4) - (-4) = 12$.

> This is my logic. Can this be proven in Abstract Algebra? Is it not an easier way to explain why $-a \times -b = +ab$?

This is essentially the same explanation given above, just with a few steps skipped over.

The key point in your explanation is that -3×-4 should be the same thing as $-(-4) - (-4) - (-4)$. The question left to answer is, why?

Everybody can accept that taking 3 times -4 is the same thing as adding -4 together three times. The question is, why does this imply that taking -3 times -4 should be the same thing as *subtracting* -4 three times?

The answer is precisely because of distributivity. -3×-4 should be the negative of 3×-4: that number which, when added to 3×-4, gives zero.

Since 3×-4 is -4 added together 3 times, what you'd need to do to it to get zero is to add 3 copies of $-(-4)$ to cancel them out. That is why, from the fact that $3 \times -4 = (-4)+(-4)+(-4)$, it follows that $(-3) \times -4 = -(-4) - (-4) - (-4)$.

The Number Zero

Asked by Estelle Shields, Sophomore, Rivier College on November 8, 1996:

> Research material needed on the number "zero". I am doing paper on this subject and am having a terrible time finding documentation.
>
> Any help you can give me is MOST appreciated!

There's not much one can say specifically about the number zero. It's just like any other whole number: it's an abstract concept measuring the size of a set. The number zero measures the size of the set with no elements in it, just like the number one measures the size of the set with a unique element in it, and so on.

Some concept of the number zero was probably in use as early as human beings first began to do arithmetic, so it's impossible to trace its origin. However, there was no *notation* for the number zero until probably somewhere between 500 and 800 A.D., and this lack of notation made it impossible for ancient mathematicians and philosophers to work with it in the same way that a modern person would.

The origin of our current notation for the number zero is also unknown, though it is presumed to have originated in India with Hindu mathematicians, somewhere between 500 and 800 A.D. Later the Hindu numerals were employed by the Arabs (and became called "Arabic numerals"), and from them spread into European mathematics.

I hope this is of some help.

Why You Can't Divide Nine By Zero

Asked by Lee Williams, student, Aldridge State High School on August 8, 1997:

> What is the answer to 9 divided by 0 on a high school level and university level.

The answer to this question is that there is no answer. By this we simply mean that there is no number which, when multiplied by 0, gives you 9.

The question "what is 9 divided by 0" is simply another way of asking the question "which number, when multiplied by 0, gives you 9?" There is no such number, and therefore no answer to the question.

This is not simply a quirk of our usual number system: in *any* number system that satisfies certain basic properties such as distributivity and associativity, 0 multiplied by anything will give 0. So if you start with any non-zero number x, there cannot be any number which when multiplied by 0 gives you x, so there can be no answer to the question "what is x divided by 0". Mathematicians say that "division by 0 is undefined", meaning there is no way to define an answer to the question in any reasonable or consistent manner.

In general number systems, "0" is defined to be the unique number which has the property that $a+0=0+a=0$ for all a in the number system. Also, for every number a there is a number $-a$ called the additive inverse of a which satisfies the property $a+(-a)=0$. Here is how to show that any number a times 0 gives you 0 (and thus that division by zero is always undefined):

$(a)(0)=(a)(0+0)$ (since $0=0+0$ by the definition of 0)

$(a)(0)=(a)(0)+(a)(0)$ (distributivity is used)

$(a)(0)+[-(a)(0)]=(a)(0)+(a)(0)+[-(a)(0)]$ (adding -(a)(0) to both sides)

$0=(a)(0)$ (since $(a)(0)+[-(a)(0)]=0$ by the definition of $-(a)(0)$)

There are some contexts in which it makes sense to talk about an "infinity" concept; see the pages on Does infinity exist? However, the above reasoning shows that you cannot simply include "infinity" in a number system and say that 9 divided by zero is infinity, at least not if you want to retain properties like distributivity which are so essential to the nature of numbers.

One final comment: the question "what is 0 divided by 0" is a little different from questions like "what is 9 divided by 0". Instead of there being no number which will work, now every number will work (every number, when multiplied by 0, gives 0). But there's still no answer to the question "what is 0 divided by 0", because this question is really asking "what is the one special and unique number which, when multiplied by 0, gives 0?" There isn't any such single, unique number, and hence there is no answer to the question.

Why is $x^0 = 1$?

Asked by Deliakos Argiris, T. E. I. (in Greece) on October 13, 1996:

> Sorry for my English.
>
> I do not understand why we agree with the axiom : $x^0 = 1$.

When b is a positive integer, a^b is defined to be the product of a multiplied by itself b times.

The question is, what is the most natural way to extend this definition to the case when $b=0$? Here are several ways to see that the definition $a^0=1$ is the only reasonable one:

1. Exponentiation satisfies the laws of exponents: $a^{b+c}=a^b a^c$. If we want this law to still be satisfied when we extend to the case $b=0$, we need to have $a^c=a^{0+c}=a^0 a^c$, and therefore we need to have $a^0=1$.

2. If a^b is b copies of the number a, all multiplied together, then a^0 should be the "empty product" with no factors multiplied together. In mathematics, the empty product is defined to be 1, because multiplying by nothing at all is the same as multiplying by 1.

3. Notice that a^b can be thought of as "start with the number 1, then multiply by a, b times." For instance, $a^2 = 1 \times a \times a$ and $a^1 = 1 \times a$. Therefore, a^0 should be just 1, not multiplied by anything else at all.

4. When a is a positive integer, yet another reason for defining $a^0 = 1$ is that a^b is the number of ways of writing (in order) b numbers, each from 1 to a. For instance, $3^2 = 9$ because there are nine different pairs of numbers each of which is in the range from 1 to 3 (they are $(1,1)$, $(1,2)$, $(1,3)$, $(2,1)$, $(2,2)$, $(2,3)$, $(3,1)$, $(3,2)$, and $(3,3)$).

   Therefore, a^0 should be the number of ways of writing no numbers, each of which is from 1 to a. There is exactly one way of doing this, namely, don't write any numbers at all!

   (This reason is more compelling if you make it more mathematically precise, using the fact that a^b is the number of functions from a b-element set B to an a-element set A, and when $b = 0$ the set B is the empty set, and there is exactly one function from the empty set into A, namely, the empty function).

The above reasons all illustrate why defining a^0 to be 1 is the only reasonable definition.

There's one other point worth mentioning: some of the reasons above are less compelling when $a = 0$. For instance, in the first reason, we need to have $a^c = a^0 a^c$, and if a is non-zero we can divide by a^c to deduce that $a^0 = 1$. However, if $a = 0$ we no longer get a reason for a^0 to be 1.

Some of the reasons are still compelling, and, especially if we are in a context where only integer exponents are being considered, we still normally define 0^0 to be 1.

However, if we define a two-variable function $f(x, y) = x^y$, then this function *does not have a well-defined limit* as $(x, y) \to (0, 0)$. We can define $0^0 = 1$ if we like, but the limit still won't exist. In other words, if A and B each approach zero, there's no guarantee as to what (if anything) A^B approaches. It need not approach our definition of 0^0.

That's why, in calculus, 0^0 is often called an *indeterminate form*. If one is working in situations where the exponent can continuously vary, it is usually better to leave 0^0 undefined to avoid making mistakes. However, if one is working in situations in which the exponent is always integral, 0^0 is usually defined to be 1.

These complications are only for 0^0. When a is nonzero, a^0 is always defined to be 1, for the reasons given above.

How To Express A Repeating Decimal Number As A Fraction

Asked by Dana Shaddad, student, Qatar International on September 22, 1997:

> How do you find a rule for expressing any recurring decimal as a fraction and such rule to be tested with examples of three digits, four digits, five digits repeating patterns.

There is a rule for converting a repeating decimal number into a fraction. Let's first of all suppose that the repeating pattern starts immediately after the decimal point, with zero before the decimal point. Let's say the repeating part is n digits long, so the number looks like $d = 0.(a_1)(a_2)\ldots(a_n)(a_1)(a_2)\ldots(a_n)\ldots$

Let A stand for the integer $(a_1)(a_2)\ldots(a_n)$ (the n-digit number whose n digits are a_1, a_2, etc.). Then the first part of of our number, $0.(a_1)(a_2)\ldots(a_n)$, is the same as $A/100\ldots0 = A/10^n$. The next part of our number is $0.00\ldots0(a_1)(a_2)\ldots(a_n)$, which is $A/10^{2n}$. The next part is $A/10^{3n}$, and so on.

This means that our number d is equal to the infinite sum $A/10^n + A/10^{2n} + A/10^{3n}$. This is a geometric series and by a well known formula (as described in another question) its value is $(A/10^n)/(1-(1/10^n)) = A/(10^n-1)$.

To see that this formula works, consider the decimal $0.142857142857\ldots$. In this case $A = 142857$, $n = 6$, and $10^n - 1 = 10^6 - 1 = 1000000 - 1 = 999999$. Then this decimal is equal to $142857/99999$. This fraction then reduces to $1/7$.

If you have a decimal where the repeating part doesn't start right after the decimal point, first multiply or divide by 10 enough times so that the repeating part does start right after the decimal point, then undo your multiplication or division at the end.

For example, suppose you have the decimal $d = 3.4575757\ldots$. Multiply by 10 to get $10d = 34.575757\ldots$. Our formula tells you that the $0.575757\ldots$ part equals $57/99$, so $10d = 34 + 57/99$, and therefore $d = 34/10 + 57/990$.

Asked by Peter Collins, student, Mill Hill School on January 2, 1998:

> How can you calculate what a fraction will be in decimal form? More specifically: How can you tell if a fraction is a recurring decimal or not, and if it is what will be the value of the decimal? (without a calculator).
>
> This is in relation to the question posted in September about writing a decimal in fractional form.

To calculate a fraction in decimal form, you perform long division. You should be familiar with this procedure from early arithmetic lessons. For example, to calculate $3/11$, you say that 11 goes into 3 zero times, with remainder 3. So the answer starts 0. Then you find that 11 goes into 30 two times, with remainder 8. So the answer starts 0.2. Next you find that 11 goes into 80 seven times, with remainder 3. So the answer starts 0.27. Now that you have a remainder 3 which you have seen before, you see that the pattern will repeat: $0.272727\ldots$.

Every fraction produces a recurring decimal. That is because there are only a finite number of possibilities for the remainder, so eventually you will encounter a remainder you have seen before, and the pattern of digits will continue from that point on.

Some fractions will give you a decimal where the recurring part is all zeros, such as $4/5 = 0.80000\ldots = 0.8$. Such decimals are called *terminating decimals.* The fractions which give you terminating decimals are those which can be written in a form where the denominator is a power of 10 (for example, $4/5 = 8/10$). Which fractions can be written this way? Precisely those fractions for which, when you write them in lowest terms (factoring out any common factors from top and bottom), have the property that the denominator is a product of 2's and/or 5's but no other prime factors. Then you can multiply top and bottom by an appropriate number of 2's and 5's to get a power of 10.

For example, $37/40$ will give you a terminating decimal, because 40 is 2 times 2 times 2 times 5. If you multiply top and bottom by two more 5's, i.e. by 25, then the denominator will become $2 \times 2 \times 2 \times 5 \times 5 \times 5 = 10 \times 10 \times 10 = 1000$ so $37/40 = (37 \times 25)/1000 = 925/1000 = 0.925$. (This is the same answer you would get if you carried out the long-division procedure above). However, $32/39$ will give you a non-terminating decimal because, even though it is in lowest terms, the denominator $39 = 3 \times 13$ has prime factors other than 2 and 5, so there is nothing you could multiply it by to get a power of ten.

Published Proof of Fermat's Last Theorem

Asked by steve bechtold, teacher, blackduck high school on November 4, 1997:

> Has a complete proof of fermat's last theorem been published? If so when was it published and has it stood up to scrutiny?

Yes, a complete proof of Fermat's Last Theorem has been published. The original proof announced by Andrew Wiles had a gap, but that gap was closed in joint work by Wiles and Richard Taylor.

The complete proof consists of two papers:

- *Modular elliptic curves and Fermat's Last Theorem*, by Andrew Wiles.
- *Ring theoretic properties of certain Hecke algebras*, by Richard Taylor and Andrew Wiles.

Both papers were published in the May 1995 issue of the journal *Annals of Mathematics*. The second paper (much shorter than the first) contains the corrected proof of the step in the argument where Wiles's original work was flawed. The first paper contains the rest of the argument. Together they constitute a complete proof of Fermat's Last Theorem, which has stood up to the scrutiny of every expert in the field. In fact, other people upon reading the proofs have come up with generalizations and simplifications of much of it.

The $n = 4$ Case of Fermat's Last Theorem

From Trevor, University at Buffalo on November 3, 1996:

> I want to prove, using mathematical induction, that there are no solutions to the equation $x^4 + y^4 = z^4$, for positive values of x, y, z.

I assume that x, y, and z are supposed to be *integer* values; otherwise, there are plenty of solutions.

Those reading this page will likely recognize this question as the $n = 4$ case of Fermat's Last Theorem. The general theorem (that, when $n > 2$, there are no positive integer solutions to the equation $x^n + y^n = z^n$), was conjectured by Fermat hundreds of years ago but remained unproved until just recently, when it was proved by Andrew Wiles. The proof is highly complex and involves some deep areas of abstract mathematics.

The $n = 4$ case, however, is relatively easy to prove and was known by Fermat. It turns out to be a little easier to prove the more general result that there is no solution in positive integers to the equation $x^4 + y^4 = z^2$ (that is, the sum of two fourth powers cannot even be a perfect square, let alone a fourth power).

This is traditionally proven using the "method of infinite descent". The key to this method is the key lemma below (I'll explain later how this fact is proven).

Key Lemma. If $x^4 + y^4 = z^2$ where x, y, and z are positive integers, then there would exist other positive integers u, v, and w with $u^4 + v^4 = w^2$, and $w < z$.

(For those unfamiliar with the terminology: a *lemma* is a small theorem whose main use is in proving another, more important, theorem).

The reason this lemma implies that solutions cannot exist is as follows. Suppose solutions did exist. Among all solutions, pick one with the smallest z value (any nonempty set of postive integers has a smallest element). But now the lemma says there would be another solution with a still smaller z value, a contradiction. Therefore, solutions cannot exist.

You can phrase this using the language of mathematical induction, but it's more awkward. You would work by induction on z. When $z = 1$ there are clearly no solutions (x and y have to be at least 1, so $z^2 = x^4 + y^4$ has to be at least 2).

Now, if it is known that there are no solutions with $z \leq N$, you can prove that there are no solutions with $z = N + 1$ either: if (x, y, z) were such a solution, then the lemma implies that there is another solution (u, v, w) with $w < z$, so $w \leq N$, contradicting the fact that we know no such solution exists.

The above two paragraphs form the basis and induction steps for a proof by induction. But it's a little cleaner to use the method of infinite descent.

Now for the hard part: a proof of the lemma. I won't fill in all the details (since you originally posted this in the *Discussion Topics* section rather than the *Questions* section, I assume you don't want me too), but I will outline the proof.

It depends on the theory of Pythagorean triples. A triple (a, b, c) of positive integers is called a *Pythagorean triple* if $a^2 + b^2 = c^2$. It is called a *fundamental Pythagorean triple* if a, b, and c have no common factor greater than 1.

For any fundamental Pythagorean triple, one of a and b is even and the other is odd. Let's use the letter a to refer to the even one, b to the odd one. Then there are positive integers m and n such that

$$\begin{aligned} a &= 2mn \\ b &= m^2 - n^2 \\ c &= m^2 + n^2 \end{aligned}$$

and m and n are relatively prime (have no common factor greater than 1).

If you don't know how to prove the above statements, feel free to post another question here.

Now, suppose $x^4 + y^4 = z^2$. Then (x^2, y^2, z) is a Pythagorean triple.

If x, y, and z have a common prime factor $p > 1$, then $(x/p)^4 + (y/p)^4 = (z/p^2)^2$ and we have our desired solution (u, v, w) with $u = x/p$, $v = y/p$, and $w = z/p^2 < z$. (You need also to explain why z/p^2 is an integer; I'll leave that to you).

On the other hand, if x, y, and z do *not* have a common prime factor, then (x^2, y^2, z) is a fundamental Pythagorean triple. One of x and y is even and the other odd; let's use x to denote the even one. Then, by the theory of Pythagorean triples, there are relatively prime positive integers m and n such that

$$\begin{aligned} x^2 &= 2mn \\ y^2 &= m^2 - n^2 \\ z &= m^2 + n^2. \end{aligned}$$

We can rewrite the second equation as $n^2 + y^2 = m^2$. We know that m and n have no common factor greater than 1, so (n, y, m) is a fundamental Pythagorean triple. We know y is odd, so n must be the even one. Therefore, appealing to the theory of Pythagorean triples once again, there are relatively prime positive integers r and s such that

$$\begin{aligned} n &= 2rs \\ y &= r^2 - s^2 \\ m &= r^2 + s^2. \end{aligned}$$

The final piece of the puzzle is the fact that if the product of two relatively prime positive integers is a perfect square, then each individually is a perfect square. From this it follows that m and $n/2$ are perfect squares, since the product is $(m)(n/2) = (2mn)/4 = (x^2)/4 = (x/2)^2$ (remember that x and n are even, so $x/2$ and $n/2$ are integers).

It also follows that r and s are perfect squares, since the product is $rs = (2rs)/2 = n/2$ which is a perfect square.

Therefore, setting $r = u^2$, $s = v^2$, $m = w^2$, the equation $m = r^2 + s^2$ becomes $u^4 + v^4 = w^2$. All that remains to complete the proof of the lemma is to show that $w < z$, which follows because $z = m^2 + n^2 = w^4 + n^2 > w^4$, so $w^4 < z$, and as w is a positive integer this implies $w < z$ as well.

If you want proofs of some of the statements I've left for you to fill in, just post a message here and we'd be glad to fill them in for you.

The Case n=3 Of Fermat's Last Theorem

Asked by Tommaso Russo on August 26, 1997:

> I've seen the proof for the $n = 4$ case of Fermat's Last Theorem
>
> Is the the $n = 3$ case similarly easy to prove? Was the proof known by Fermat?

The proof that there are no integers X, Y, and Z which satisfy the equation $X^n + Y^n = Z^n$ when $n = 3$ is similar to the proof in the case where $n = 4$ with the exception of a crucial lemma. The statement of the lemma is stated as follows:

If x, y, and z are integers such that $x^2 + 3y^2 = z^3$ and x and y are relatively prime then there exist integers a and b such that $x = a^3 - 9ab^2$ and $y = 3a^2b - 3b^3$.

The proof of this lemma hinges on some material which is typically covered in an advanced undergraduate or an introductory graduate abstract algebra course. Those who are interested in more reading on the subject and who have enough background in mathematics can find this lemma (together with hints) as exercise 4.6 in Daniel Flath's *Introduction To Number Theory* (see also exercises 7.6 and 7.8 for more information on cases $n = 3$ and $n = 4$ of Fermat's Last Theorem).

A Question from the IMO

Asked by Andrew Mnih, Father Redmond High School on July 31, 1996:

> I have been working on question #4 from this years IMO paper:
>
> 4. The positive integers a and b are such that the numbers $15a + 16b$ and $16a - 15b$ are both squares of positive integers. Find the least possible value that can be taken by the minimum of these two squares.
>
> Can anyone help me solve this?

(Note: the "IMO" is the "International Mathematical Olympiad")

The conditions given can be translated into $15a + 16b = c^2$ and $16a - 15b = d^2$ where a, b, c, and d are positive integers. You are asked to find what is the smallest that the smaller of c^2 and d^2 can be.

Let's try and translate all of this information into conditions on c and d. We know they have to be positive integers. The remaining conditions are that a and b have to be positive integers. So, let's solve the above system of equations to express a and b in terms of c and d.

If you do this, and use the fact that $15^2 + 16^2 = 21^2$, you get $21^2 a = 15c^2 + 16d^2$. This is guaranteed to be positive, so the only other thing necessary is that a be an integer, i.e.

> $15c^2 + 16d^2$ must be a multiple of 21^2. (**Condition A**)

Solving for b gives $21^2 b = 16c^2 - 15d^2$. This means that c^2 must be greater than $(15/16)d^2$, and that

> $16c^2 - 15d^2$ must be a multiple of 21^2. (**Condition B**)

Summarizing: the only conditions on the squares c^2 and d^2 are that they're positive, satisfy the inequality $c^2 > (15/16)d^2$, and satisfy conditions A and B.

Whenever you have a condition like A or B, it's useful to consider prime factors. The prime factors of 21^2 are 3 and 7. Let's work with 3 first. Since 3 divides 21^2 and 3 also divides $15c^2$, condition A tells us that 3 must also divide $16d^2$ and hence must divide d.

Now I'm going to stop giving you the solution and see if you can work out the rest for yourself. You should be able to show that Conditions A and B together are equivalent to a very nice and simple condition on c and d, so that you can tell exactly which numbers are possible values for c and d, and hence from that determine what the smallest possible value is for one of them.

Post another message here if you get stuck further along, or if something I said here wasn't clear.

How To Find The Least Common Multiple

Asked by Ken Sharp on August 3, 1997:

> What is the lowest number which has every number from 1 to 10 as a factor? Is it 2520? This number is divisible by every number between 1 and 10, but is it the lowest number. Does 'the' lowest number have any special name?

This lowest number is often called the *least common multiple* (or LCM for short). To compute it you must first write out the prime factorization of each of the numbers in question: $2 = 2$, $3 = 3$, $4 = 2 \times 2$, $5 = 5$, $6 = 2 \times 3$, $7 = 7$, $8 = 2 \times 2 \times 2$, $9 = 3 \times 3$, and $10 = 2 \times 5$. Note that we left out 1 as it has no bearing on the value of the LCM.

For each prime listed in one of the above factorizations, find the factorization where it occurs the most. For instance 2 appears most frequently in the factorization of 8, where it appears three times. The most times 3 appears is twice (in 9), and the most 5 and 7 appear is once.

The LCM is the product of all of these primes, each one repeated the number of times it occurs in the factorization where it appears most frequently.

Therefore, the LCM of the numbers from 1 through 10 will be the product of the number 2 (occurring three times), the number 3 (occurring twice), and the numbers 5 and 7 (occuring once). In other words, the LCM is $(2^3)(3^2)(5)(7) = 2520$.

The reason this is a *multiple* of every number between 1 and 10 is that any number from 1 to 10 can be written as a product of primes, each occurring at most as often as it does in 2520. Therefore, 2520 is a multiple of that number (it is that number times whatever primes are left over).

For example, $6 = (2)(3)$, and we can take one 2 and one 3 from 2520 to get $2520 = [(2)(3)][(2^2)(3)(5)(7)] = (6)(420)$. The same works for any other number from 1 to 10.

The reason that this is the *smallest* multiple of every number between 1 and 10 is that, in order for one number to divide a second, every prime factor of the first number must divide the second number at least as many times as it divides the first. Thus *any* common multiple of 1 through 10 must have 2^3 as a factor (since 8 does), must have 3^2 as a factor (since 9 does), and must have 5 and 7 as factors. Therefore any common multiple of 1 through 10 must have $(2^3)(3^2)(5)(7) = 2520$ as a factor, proving that 2520 is the *lowest* (or "least") common multiple.

The above arguments also work for finding the LCM of any collection of numbers, not just 1 through 10.

A Geometric Proof That The Square Root Of Two Is Irrational

Asked by Robert Second on August 28, 1997:

> I heard somewhere that there is a proof that root 2 is irrational by geometric means. Does anyone know this?

The geometric proof is a somewhat more awkward version of the proof that is most commonly given for the irrationality of $\sqrt{2}$. Its significance is mostly historical—it was the first known proof, discovered by the Pythagoreans. This proof was taken from Euclid's *Elements*, volume III.

Let $ABCD$ be a square (with diagonal AC) and consider the ratio $AC : AB$. Suppose for contradiction that $AC : AB = n : m$ for two positive integers m and n which have no common divisor. It can be seen (either by the Pythagorean Theorem or by comparing areas) that $AC^2 = 2AB^2$ and thus $n^2 = 2m^2$. Note here that $\sqrt{2} = AC/AB$. It follows from our assumptions that n is even and m is odd. Since n is even, let $n = 2k$. Then $4k^2 = 2m^2$ and $m^2 = 2k^2$. But this implies that m is also even, a contradiction. Thus $AC/AB = \sqrt{2}$ is not a rational number.

Asked by Robert Second, student, Greendale on October 3, 1997:

> I don't understand how either proof of root 2 is an irrational number (the geometric method and the contradiction method) works. How does proving that numbers are even prove that root 2 is irrational??
>
> Thanks.

Any rational number can be expressed as a fraction in lowest terms, that is, in the form a/b where a and b have no common factors. In particular, this means that any rational number can be expressed in the form a/b where a and b aren't both even.

To put it another way: suppose you start with some fraction a/b. It's certainly possible that a and b might both be even, but if they are, you could divide both of them by two and the fraction will still represent the same rational number. You can keep on doing this as long as they are both even. Since no integer can be divided by 2 infinitely often and remain an integer, this process must stop sometime, and you will end up with a fraction where the numerator and denominator are not both even.

For example, if you start with the fraction 8/12, after twice dividing top and bottom by 2 you end up with 2/3 and 3 is odd.

Therefore, if the square root of 2 were rational, you would be able to write it in the form a/b where a and b are not both even. However, the proof shows that this is not possible, and therefore the square root of 2 is not rational.

Asked by Bryan Low, teacher, San Leandro High School on January 3, 1998:

> Why is the square root of of 2 irrational? Is there another proof besides the geometric one founded by the Pythagoreans? I have read that one on your site. I don't believe my student remeber much geometry to understand that irrational proof posted on your site.
>
> Any help will be appreciated!!!
>
> Bryan Low

You don't need any geometry for the proof; it was just originally phrased that way because the Pythagoreans were thinking about geometry at the time.

The essence of the proof is this. If the square root of two were a rational number, you could write it as a fraction a/b in lowest terms, where a and b were integers, not both even. (If a and b were both even, the fraction wouldn't be in lowest terms; you could divide top and bottom by 2, and keep doing this until one of them stops being even).

That means that, if the square root of two were a rational number, it would be possible to find two integers a and b, not both even, such that $a/b = \sqrt{2}$. But this is impossible, because the equation can be written as $a^2/b^2 = 2$, so $a^2 = 2b^2$, so a is even. That means you can write $a = 2k$ where k is an integer. Now the equation $a^2 = 2b^2$ becomes $4k^2 = 2b^2$, so $b^2 = 2k^2$, so b would have to be even as well.

Therefore, because it is impossible to find two integers a and b with the property that a and b are not both even and $a/b = \sqrt{2}$, the square root of two cannot be a rational number.

Which U.S. President Re-Proved the Pythagorean Theorem?

Asked by Autumn Lambert, student, Warner Robins High on March 12, 1997:

> Which President designed a proof for the pythagorean theorem?

I believe that the U.S. President Ulysses S. Grant came up with a new proof of the Pythagorean Theorem. (I will attempt to verify that information to make sure).

Why Arithmetic and Geometric Sequences are Called What They Are

Asked by Flavia Fayet, student, Vaughan Secondary School on December 18, 1996:

> Hello,
>
> My question is not an actual problem, but it's been puzzling me and I haven't found an accurate answer yet. Maybe you could help me.
>
> Why are arithmetic sequences actually called arithmetic? What about them makes it arithmetic? What about geometric sequences. . . why aren't they called arithmetic? Don't they contain arithmetic in them?

> Also why is some sequences both arithmetic and geometric, like 2,2,2,2,2,2,...? The obvious answer is it that one has a common difference and the other a common ratio, but it's not completly clear to me. I believe that arithmetic sequences are called that because of thier similarity to arithmetic progressions, but it doesn't completly explain to me the real concept about it.
>
> Please reply A.S.A.P. Thank you for your time in answering my question and hope it doesn't trouble you!! :-)
>
> Yours Truly,
>
> Flavia Fayet

Geometric progressions have been found on Babylonian tablets dating back to 2100 BC. Arithmetic progressions were first found in the Ahmes Papyrus which is dated at 1550 BC. The names for these notions, however, seem to have taken considerably longer. In some cases there was no standard for how to refer to them (even the term progression was not necessarily a standard).

The closest I can come to the reasoning behind the names is that each term in a geometric (arithmetic) sequence is the geometric (arithmetic) mean of it's successor and predessor. The rationale behind the names of these means is a bit more clear: if we view the quantities A and B as the lengths of the sides of a rectangle, then the geometric mean $\sqrt{AB}$ is the length of the sides of a square having the same area as this rectangle. This was viewed in those days as a very *geometric* problem: finding the dimensions of a square having the same area as a given figure (in this case, rectangle).

Although the arithmetic mean $(A+B)/2$ can also be interpreted geometrically (it is the length of the sides of a square having the same *perimeter* as the rectangle), lengths were viewed more as arithmetical concepts (because it's easy to handle lengths by ordinary addition and subtraction, without having to think about two-dimensional concepts such as area).

You are, of course, perfectly right when you say that both concepts involve arithmetic. It is also true that both concepts can be interpreted geometrically. Nevertheless, in ancient times one was viewed much more geometrically than the other, hence the names. I hope the above paragraphs shed some light on why that was so.

Origin of the Notation for Slope

Asked by Jessica, Marquette Senior High School on Friday Mar 22, 1996:

> In the formula $y = mx + b$, why is m used for the slope? What does it stand for?

This question still has us stumped! Why and when did it become customary to use the letter m for slope?

If anybody knows the answer, please post a follow-up message using the form below. Here are a few comments that have been posted so far:

Follow-up comment by Dolores Greenberg (High School Librarian), La Habra High School:

> I have a math teacher who has posed the question about the notation "M" for slope. I have hunted everywhere without success except to find that slope stands for "M"aximum rate of change of the function. Could this be the "M"?

It's possible, although I doubt it because slope is not really anything to do with "maximum" rate of change of a function, rather, just plain old rate of change.

However, I haven't been able to come up with anything more conclusive myself. It is most likely from some Latin word, but I am not sure which; another possibility is that it may be from the French *montrer* for "to climb". I'm still trying to find out something more definitive, and I'll keep you all posted!

Follow-up comment by Dolores Greenberg (High School Librarian), La Habra High School:

> We are still working on it at this end. Our ideas to date: *montagne* Fr. for mountain, Magnitude method from siting artillery, related to M in the Greek alphabet somehow.

Why We Use "Argument" In Describing Complex Angles

Asked by Pam Deeker, teacher, St. Bede's College on July 24, 1997:

> Where did the name "Argument" come from to name the angle when complex numbers are written in polar form?

To the best of our knowledge, the word argument was originally used by astronomers when referring to certain angles associated with orbits. The argument of pericenter, for example, is defined as follows. There is a line through space which is fixed by convention as a reference. If one body orbits another (for instance the Moon orbiting around the Earth) the argument of pericenter associated with this orbit is the angle in the plane of the orbit between the reference line and the closest point of the orbit to the focus. The angle described above is measured from the focus of the orbit and in the direction which the object orbits.

Measuring the argument of a complex number is similar. The reference axis can be thought of as the x-axis. The location of the complex number is analogous to the closest point of the orbiting object to the focus or origin.

I do not know why "argument" was first used to mean angle or arc in astronomy. The earliest citation given in the Oxford English Dictionary is from Chaucer, circa 1391: "To knowe the mene mote and the argumentis of any planete" (*Astrol.* xliv. 54).

The Origin Of The Word Quadratic

Asked by Gab Fredman, student, Columbia High School on September 5, 1997:

> I have a pretty random question about the derivation of the word "quadratic. I naturally think of four when I hear the word, but in actuality, it does not mean that at all. Do you have any clue as to how the word "quadratic" got its name?

To the best of our knowledge the origin of the term "quadratic" is Latin. It is derived from *quadratus* which is the past participle of *quadrare* which means "to make square." From this it is clear that part of the word is connected to the Latin word for "four," though not a way which one might expect: it refers to squaring, and a square is a regular four-sided figure.

The Origin Of Geometry

Asked by Cristy Smith, student, Valliant High School on October 23, 1997:

> Where did geometry originate?

This is a difficult question to answer since geometry has been around in one form or another as long as there has been written history. The Egyptians were making use of geometry to build pyramids even in prehistoric times. It was the Greeks, however who first began to rigorously study geometry and try to prove facts about it. Euclid was perhaps the most notable geometer of ancient mathematics and it was he who first axiomatized the subject (carefully defined the concepts crucial to geometry).

Origin of Orders of Operations

Asked by scott states, teacher, elida high school on November 4, 1997:

> Who came up with the standard order of operations? When was it originated?

I'm going to assume you're referring to the convention that an expression like $a+3b$ is understood to mean $a + (3 \times b)$ and not $(a + 3) \times b$.

This convention is the natural one because multiplication distributes over addition: that is, any product of sums can be re-written as a sum of products (for instance, $(2+3)(4) = (2)(4)+(3)(4)$), but not every sum of products can be factored as a product of sums.

Therefore, every algebraic expression involving sums and products can always be written as a sum of products. This, then, is the most common, convenient, and natural form in which to write things. Consequently, it makes sense for this to be the form we assume when we omit parentheses in an expression involving sums and products.

That is the origin for the convention that multiplication (and division) have "higher precedence" than addition and subtraction. I do not imagine it originated with any specific person, but was the natural outcome of people wanting to be able to write expressions that were not cluttered with parentheses everywhere, and choosing to interpret such un-parenthesized expressions in the form that is the most universally applicable to all expressions: as a sum of products (for instance, the "3 + (4 times 5)" interpretation of "3 + 4 times 5") rather than as a product of sums (the "(3 + 4) times 5" interpretation), a form which occurs only in factorizable expressions.

How To Code Information With Error Detection

Asked by David Dudley on November 25, 1997:

> Please, Please could you give me an explanation of these number systems:-
>
> Two out of five code
>
> Excess3 (XS3)

We believe that "two out of five code" is a special case of what is known more generally as a Hamming code. In this case pieces of information are encoded as a string of 5 bits (0's or 1's) in such a way that errors of 1 bit or less can be detected and corrected.

To do this, we pick a collection of 5-bit binary strings such that any two strings in the collection differ in more than 2 bits. We also pick the collection in such a way that if we have any string of 5 bits, there is exactly one string in our collection which differs from it in 2 or fewer bits. Now we record encoding/decoding scheme which tells us which letters/numbers correspond to which strings. One entry in our decoding sheet might by "A = 00010," for instance.

To transfer data using this coding scheme, we take our information and write it in terms of 5-bit strings according to our encoding table. The data is then transfered or stored and recovered. To decode a 5-bit string, we find the unique string in our collection which differs from it in at most 2 bits. We then use the decoding table to convert this string into readable information.

Of course if no errors have occurred, all the strings which are received will actually appear on the list. In the event of a minor error (1 bit changed), the damaged string is still closest to the original string. Note that an error of 2 bits can also be detected, though it may not be properly corrected.

In general, if we want to use n binary digits to code numbers and be able to detect errors of size at most k (being able to correct errors of size at most $k/2$) then we search for a collection of n-bit sequences such that any two of them differ in more than k bits. We also require that any n-bit string differ from exactly one string in the collection by k or fewer bits. The encoding and decoding method is the same as described above.

We remark here that a true Hamming code may have to meet a slightly more strict (and more complicated) criterion. The details have be eliminated for simplicity.

We are unfamiliar with "Excess3."

How To Build A Parabolic Dish

Asked by an anonymous poster on October 5, 1997:

> How do you build a Parabolic dish?

While it is difficult to build a parabolic shape with a large amount of precision (in the absence of expensive machinery), there are a few tricks which allow anyone to build a crude parabolic dish.

One method is to cut out a two dimensional parabolic shape from some rigid material and then use it to help you model some three dimensional shape (out of aluminum foil or chicken wire, for instance). The focus of the parabola $y = ax^2$ is located at the point $(0, p)$ where $p = 1/4a^2$.

Although it is more difficult to implement, nature provides us with a far more accurate way of obtaining a parabolic shape. If a body of water is in a rotating container (and the liquid and the container are rotating together at the same speed), the surface of the water takes the shape of a near perfect parabolic shape. The focus of the parabolic shape depends on the rate of the rotation and also on the density of the liquid (if you were using something other than water). This approach is probably more interesting for those interested in the optics of parabolic shapes.

How To Compute Standings In Baseball

Asked by David Tobey on October 6, 1997:

> In sports statistics, namely baseball, how do you calculate the number of games that a second place team is behind the first place team? This is easy if both teams have played the same number of games, but if they have not played the same amount, it isn't so clear. This sounds like a fascinating mathematical question for a math fan and sports buff. I have come up with a few scenarios, but never have seen a definitive formula.

> I apologize if this is not the correct forum for this, but I haven't found anything yet that is. Please let me know where I can find this answer elsewhere. Your web pages seem to be a wealth of statistical and mathematical knowledge. I need to solve this dilemma for a league that I am administering.
>
> Thanks.

The standard way to interpret the number of games one team is behind another is as the number of games in which that team would have to beat the other team in order for their standings to become equal.

This number is always an integer when the teams have played the same number of games. However, when the number of games played is unequal, it's convenient to consider fractions of games. For instance, if team I has won 9 and lost 10, while team II has won 10 and lost 10, we say that team I is half a game behind team II (because if team I played and beat team II in one full game, team I then be as much ahead of team II as it used to be behind).

One way to interpret this mathematically is the following. Suppose that Team I has won w games and lost l games and Team II has won W games and lost L games. Their winning averages (wins per games played) are $w/(w+l)$ and $W/(W+L)$ respectively.

We say that Team I is n games behind Team II if adding n more wins to Team I and n more losses to Team II would cause their winning averages to become equal; in other words, if $(w+n)/(w+l+n) = W/(W+L+n)$. Solving for n we get that $n = -(w+L)/2 + \sqrt{(w+L)^2 - 4wL + 4Wl}/2$. Admittedly this formula looks complex, though it gets directly to the point by setting the winning averages of the two teams to be equal after n Team I versus Team II games are won by Team I.

The output of the formula will also need to be rounded to the nearest half game to make the statistics "look nice" and be consistent with the usual practice. For instance, if team I has a 10-and-10 record while team II has a 10-and-9 record, the above formula gives $n = 0.51249\ldots$ and we would normally consider the number of games behind to be 1/2.

Regular Withdrawals on Compound Interest

Asked by An anonymous poster on July 18, 1997:

> If a person has \$150,000 and it compounds at, let's say, eight per-cent per year, and that person (while all this is taking place) draws \$1,200 monthly from the fund, how long will the fund last until it is exhausted??? Is there a formula I can use to determine the time/amounts based on compounded interest etc.??? For example, if it was \$175,000 at 9 per-cent... how many years when drawing out \$1,300 per month?? I suspect you have a simple formula or way of doing it????? Appreciate it if you could e mail me your suggestion (short of taking a math class.) Thanks.

Your question is actually the same as the principles used in mortgage calculation. Here is how to derive the formulas:

Suppose the monthly interest rate is I, and that an amount W is withdrawn each month. What this means is that, during each month, the balance gets multiplied by $1+I$ (it becomes the original amount plus the interest, which is I times the original amount) and then has W subtracted from it.

Therefore, if $B(0)$ is the starting balance, the balance after one month will be

$$B(1) = (1+I)B(0) - W.$$

For convenience, let J denote $1+I$, so that $B(1) = JB(0) - W$. After two months the balance will be

$$B(2) = JB(1) - W = J(JB(0) - W) - W = J^2B(0) - W - JW$$

and in general, after n months the balance will be

$$B(n) = J^nB(0) - W - JW - J^2W - \ldots - J^{n-1}W.$$

There's a convenient formula for the sum $1+J+\ldots+J^{n-1}$: it is $(J^n-1)/(J-1) = (J^n-1)/I$. Therefore, abbreviating the starting balance $B(0)$ to just "B",

$$B(n) = J^nB - (J^n - 1)W/I.$$

Your question is asking how large n has to be before the balance drops to zero; in other words, you want to solve for n in the equation

$$0 = J^nB - (J^n - 1)W/I.$$

Some basic algebra lets you rewrite this equation as $J^n = W/(W - BI)$; taking logarithms of both sides gives

$$n\log(J) = \log(W) - \log(W - BI)$$

so

$$n = \log(W)/\log(J) - \log(W - BI)/\log(J).$$

The only other thing you need to know before being able to solve your problems is what the monthly interest rate is. The monthly interest rate I and annual rate A are related by $(1+I)^{12} = 1+A$ (because your balance is multiplied by $1+I$ each month and hence is multiplied by $(1+I)^{12}$ each year, but we also know that each year it is multiplied by $1+A$, so these two factors must be equal).

Therefore, $J = 1 + I = (1 + A)^{1/12}$.

In your first example, $A = 0.08$ (8 percent), so $J = 1.08^{1/12} = 1.0064\ldots$ and I is roughly 0.0064. You have $W = 1200$ and $B = 150000$. Plugging these into the above formula gives

$$n = 254.30\ldots$$

so the account would be exhausted during the 255th month (during the 22nd year).

In your second example, $A = 0.09$, $W = 1300$, $B = 175000$, giving $n = 489.28\ldots$ so the account would be exhausted during the 41st year.

These formulas are usually used in mortgage calculations: the amount you owe is increased by the interest accruing on it, but is reduced by each of your monthly payments W (just as, in your example, your bank account increases by the interest earned but is recuced by your monthly withdrawals). In mortgages, however, n is known and W is what needs to be calculated. For example, in a 25-year mortgage at 8% on \$100,000, the bank needs to calculate a monthly payment amount W which will reduce the balance to zero after $n = 300$ months. So the equation

$$0 = J^nB - (J^n - 1)W/I$$

needs to be solved for W. The solution is

$$W = J^nBI/(J^n - 1).$$

In this example, $n = 300$, $B = 100,000$, $J = 1.0064\ldots$ and $I = 0.0064\ldots$, yielding $W = 753.415\ldots$. Your monthly payments on such a mortgage would be \$753.42.

Applications of the Geometric Mean

Asked by Senthil Manick on May 22, 1997:

> When would one use the geometric mean as opposed to arithmetic mean? What is the use of the geometric mean in general?

The arithmetic mean is relevant any time several quantities add together to produce a total. The arithmetic mean answers the question, "if all the quantities had the same value, what would that value have to be in order to achieve the same total?"

In the same way, the geometric mean is relevant any time several quantities multiply together to produce a product. The geometric mean answers the question, "if all the quantities had the same value, what would that value have to be in order to achieve the same product?"

For example, suppose you have an investment which earns 10% the first year, 50% the second year, and 30% the third year. What is its average rate of return? It is *not* the arithmetic mean, because what these numbers mean is that on the first year your investment was *multiplied* (not added to) by 1.10, on the second year it was multiplied by 1.60, and the third year it was multiplied by 1.20. The relevant quantity is the geometric mean of these three numbers.

The question about finding the average rate of return can be rephrased as: "by what constant factor would your investment need to be multiplied by each year in order to achieve the same effect as multiplying by 1.10 one year, 1.60 the next, and 1.20 the third?" The answer is the *geometric mean* $(1.10 \times 1.60 \times 1.20)^{1/3}$. If you calculate this geometric mean you get approximately 1.283, so the average rate of return is about 28% (not 30% which is what the arithmetic mean of 10%, 60%, and 20% would give you).

Any time you have a number of factors contributing to a product, and you want to find the "average" factor, the answer is the geometric mean. The example of interest rates is probably the application most used in everyday life.

Here are some basic mathematical facts about the arithmetic and geometric mean:

Suppose that we have two quantities, A and B. Taking their arithmetic mean we get the number $(A+B)/2$ which can be interpreted in a number of ways. One interpretation (probably the most common) is that this quantity is the midpoint of the two numbers viewed as points on a line.

Now suppose that we have a rectangle with sides of lengths A and B. The arithmetic mean can also be interpreted as the length of the sides of a square whose perimeter is the same as our rectangle. Similarly, the geometric mean $\sqrt{AB}$ is the length of the sides of a square which has the same area as our rectangle.

It is known that the geometric mean is always less than or equal to the arithmetic mean (equality holding only when $A = B$). The proof of this is quite short and follows from the fact that $(\sqrt{A} - \sqrt{B})^2)$ is always a non-negative number. This inequality can be surprisingly powerful though and comes up from time to time in the proofs of theorems in calculus.

Asked by G. Ellis, student, Southeast Bulloch High on January 16, 1997:

> Could you give the formula for the geometric mean for a series of numbers if I am trying to get the compound annual growth rate for a series of number that include negative numbers?

In general, you can only take the geometric mean of positive numbers. The geometric mean of numbers $a_1, a_2, \ldots, a_n$ is the nth root of the product $a_1 a_2 \cdots a_n$.

In your example, you *are* taking the mean of positive numbers. For example, if you're looking at an investment that increases by 10% one year and decreases by 20% the next, the simple rates of change are 10% and −20%, but that's not what you're taking the geometric mean of.

At the end of the first year you have 1.1 times what you started with (the original plus another tenth of it). At the end of the second year you have 0.8 times what you started the second year with (the original minus one fifth of it). So, the numbers you are taking the geometric mean of are 1.1 and 0.8. This mean is approximately 0.938.

This means that, on average, your investment is being multiplied by 0.938 (= 93.8%) each year, a 6.2% loss.

So, the compound anual growth rate is (approximately) −6.2%.

Asked by Paul van Esbroeck on October 5, 1997:

(Question abridged from original posting)

> I am presently engaged in a dispute with a bank concerning a Stock Market Tracker GIC. I have found significantly different definitions for the terms "average percentage growth" and "average percentage growth rate", and in many instances I cannot distinguish between the use of the words "growth" vs. "growth rate" in the literature. It was in investigating growth that I came to your question page about the geometric mean.
>
> While I, like most people I asked, originally understood these GICs to be guaranteeing a rate of return equal to the rate of growth of the index, the bank has a different interpretation.
>
> Let's say a stock market index starts at 1000, and at the end of 1 month is 1010, at the end of 2 months is 1020, and so on, ending up at 1120 after one full year.
>
> The bank seems to be considering the growth by month end for each month (which is 10 for the first month, 20 for the second, and so on, with this growth being 120 by the end of the twelfth month), then averaging these growths, obtaining an "average growth" of 60 and an average percentage growth of 60/1000 or 6%.
>
> For other definitions there seems to be quite a bit of agreement that:
>
> - The real growth was 120/1000 = 12%.
> - The average percentage growth rate is 12% per year.
> - The annual compound growth rate is 12%.
>
> It seems to me that it is incorrect to average growth as done by the bank. Since an average should be of equal periods, in the banks formula we average growth for the first month with growth for the first 11 months. Since the growth was 10 index units each month, should the average growth not be 10 units per month, but how is this different from the average growth rate?
>
> What if not growth, do you call the plot of (Index value minus Index value at some starting time) ? What if any, is the correct distinction between growth and growth rate ? Do you find the term "average percentage growth" as used by the bank problematic?

First let's try to get the words sorted out. Then we can address the situation you describe.

The word "growth" is often used quite loosely to mean any of the above concepts. The best and most correct definition of growth would be a quantity that is associated to a particular period in time, and describes the index value at the end of the period minus the index value at the beginning of the period. In your example, the growth for the period consisting of the first month was 10, and the growth for the period consisting of the first 12 months was 120.

"Growth rate" describes growth per unit time. It may vary with time. The "average growth rate" over a period is the growth divided by the length of the period. If the growth rate is constant over a period, then the average growth rate over the period will be the same as that constant value.

In your example the growth rate was a constant 10 units per month. The average growth rate over every period was also 10 units per month. For example, the average growth rate during the first month was 10 units per 1 month = 10 units per month. The average growth rate during the entire year was 120 units per 12 months = 10 units per month.

These concepts are not relevant to typical investements, stock market indices, etc., because the growth rate tends to be proportional to the index value. After all, if a 10 dollar investment grew by 1 dollar over a year (an average growth rate of 1 dollar per year), you'd expect a 10,000 dollar investment to grow by 1000 dollars per year: a very different growth rate!

So instead, the relevant concepts are *percentage* growth and growth rates. The percentage growth over a period is the ratio of the growth to the starting value. In your example, the percentage growth during the first month was 1%. The percentage growth during the second month was about 0.99% (10/1010). The percentage growth during the entire year was 12% (120/1000).

The percentage growth rate is the ratio of the growth rate to the current value. Thus, if an index value is 1000 and it is growing at 10 units per month, it is experiencing a percentage growth rate of 1% per month. If its value is 1010 and it is growing at 10 units per month, its percentage growth rate is about 0.99% per month. And so on. In your exaple, by the end of the year it is still growing at 10 units per month and its value is 1120, so the percentage growth rate by the end of the year has dropped to about 0.89%.

The only tricky thing about percentage growth rate is in changing units: a percentage growth rate of 1% per month is not the same thing as 12% per year. If an index were growing at a constant percentage growth rate of 1% per month, after one month it would be 1.01 times its starting value, after two months it would be 1.01^2 times its starting value, and after 12 months it would be $1.01^{12} = 1.1268\ldots$ times its starting value.

Therefore, a percentage growth rate of 1% per month is the same as a percentage growth rate of about 12.7% per year.

If the percentage growth is G over a period of T time units, then the average percentage growth rate over that period is $(1+G)^{1/T} - 1$.

To figure out the average percentage growth rate of your example over the entire year: if your units are years, then $T = 1$. $G = 12\%$, so the average percentage growth rate is $(1+G)^1 - 1 = G = 12\%$ per year. If your units are months, the average percentage growth rate is $(1+G)^{1/12} - 1$ which is about 0.949% per month. These are each saying the same thing; an average percentage growth rate of 12% per year is the same as an average percentage growth rate of about 0.949% per month.

Often, in investment circles, when people refer to "growth" or "growth rate" they are meaning percentage growth and percentage growth rate.

Now, on to your particular situation.

As you point out, it is not particularly meaningful to average the growths over time periods of wildly differing lengths. So, to average the growth during the first month with the growth during the first 12 months is not a reasonable thing to do.

However, what *is* reasonable to do is to average the growths over overlapping time periods of similar lengths, to smooth out fluctuations in the index. For example, let's suppose the end of the year happened to be a really bad day on the stock market, and instead of rising to 1120 the index fell back to 1000, jumping back up to 1120 the next day. Would you really want to say that there was no growth at all during the year, just because the day you picked to evaluate the index happened to be a bad one? I think not.

So, what is commonly done is to average the index value over a certain time period. For example, while what you describe is unreasonable, it could be perfectly reasonable to say that the GIC's rate of return from 1996 to 1997 will be the percentage growth of the index's *average* 1996 value to its *average* 1997 value. That way, fluctuations due to the particular day the index is measured will be smoothed out.

Now, if the index were a constant 1000 during all of 1996, and rose to 1120 in the manner you describe during 1997, you would unfortunately only be getting a 6% return from average 1996 value (1000) to average 1997 value (1060). But that's because some of the 1996 behaviour is being factored in as well.

Another thing that could be reasonable is to average the percentage growths over different 12-month periods ending in the same year. For example, you could take the January 1996-January 1997 percentage growth, the February 1996-February 1997 percentage growth, and so on, and average them. This is a reasonable sort of average to take, since the percentage growths are being averaged over equal but overlapping periods.

These sorts of averages have the advantage of smoothing out wild fluctuations in the index. For example, if the index started 1996 at 1000 and rose at 10 units per month steadily, ending 1997 at almost 1240 but plummeting down to 1150 on the last day of 1997 due to a temporary correction in the markets, you would still get the benefit of the growths. Even though the December 31, 1996 to December 31, 1997 growth was from 1120 to 1150, just 2.7% on the entire year, the November 30, 1996 to November 30, 1997 growth was from 1110 to 1230, about 11%. So, by including these in the average, rather than just looking at the average percentage growth over a single year-long period, you get a much less volatile measure of performance.

My suspicion is that the GIC you are concerned about probably employs some sort of legitimate averaging such as the above, and that the bank has miscommunicated the calculation method to you.

Another possibility is that the bank may be measuring index value from start of investment to the average index value during the year preceding the end of investment. That means that during the first year of the investment you have the unreasonable calculation you described, but over the long term those effects are negligble. For instance, the 10-year return might be the average of the index return over the twelve periods all beginning on the date of invesment and ending on the twelve month-end dates during that 10th year. True, this is still averaging growth over periods of differing length, but the lengths range from 9 years to 10 years and this is much less of a difference than your example where they range from 1 month to 12 months!

I will not venture to speculate further on what methods the bank may or may not be using for the GIC you have in mind. However, I hope this clears up for you some of the mathematical ideas involved.

Scientific Notation in Everyday Life

Asked by Johnathan Marshall and Christina Dimingko, students, Brookville on December 11, 1996:

> What is the use of scientific notation in every day life?

Scientific notation is needed any time you need to express a number that is very big or very small. Suppose for example you wanted to figure out how many drops of water were in a river 12 km long, 270 m wide, and 38 m deep (assuming one drop is one millilitre). It's much more compact and meaningful to write the answer as roughly 1.23×10^{14} than it is to write 123120000000000.

For one thing, the scientific notation is easier to read, and makes it much easier to tell at a glance what the order of magnitude is (rather than counting zeros).

For another, most of the digits in 123120000000000 are completely meaningless (unless your measurements were very precise). For instance, if the exact river length were really 12.123123 km (we just measured it to the nearest kilometre), then correct number of drops would be 124383242000000, and after the first three digits our result of 123120000000000 is quite inaccurate. So it's better to use a notation (like scientific notation) in which you can suppress the inaccurate digits.

Followup question by an anonymous poster on February 11, 1997:

> Who created scientific notation? What are the uses for it in the work field?

Scientific notation was not "created", in the sense of someone coming up with something new. The fact that 3×10^4 happens to equal 30000 is a mathematical truth, not a creation.

The question becomes, though, when did it become commonplace to write the first form instead of the second form. (It would be sort of like people starting to write $2+3$ whenever they meant 5; that's not creating something new, merely saying something in a different way).

I do not know who first used scientific notation. The concept would be very old; you'd have to dig back to the first time someone thought of describing 10000000000 as "a one followed by ten zeros", realized that's the same as 10^{10}, and wrote it that way (in whatever notation they were used to using for exponents).

The modern notation for exponents (writing them raised at a higher level) originated with Descartes in 1637, so you would never have seen an expression like 3×10^4 before then. Sometime between then and the present it became common to write large and small numbers that way, as well as numbers where it's important to convey an indication of the precision of a measurement; I do not know when it became common practice or who started doing it, but I will see if I can find out. It most likely occurred during the 1800's and 1900's when scientists were developing their understanding of the astronomical universe (involving really huge numbers to describe distances), and of the world of subatomic particules (involving really small numbers).

I don't know that I can say in answer to the question "what are the uses for it in the work field", other than what I've already said in answer to the previous question on this page: it would be needed any time you are dealing with numbers that are very large or very small, and any time you make a measurement of something and want to write the number in a way that gives an indication of its precision.

For example, if you're an engineer and you want to record the pressure on a supporting beam of a bridge, and you measure it as 500034 but your instrument is only precise to ± 600, you would not want to write "500034" because you really have no way of knowing, based on your measurement, what the last few digits are. On the other hand, you wouldn't want to just round it to 500000, because that doesn't convey the fact that you *do* precisely know the first few digits! Scientific notation (5.00×10^5) is the perfect way to express the number *and* give an idea of how precise it is.

So, the answer to your question is, just pick any field in which people deal with large and small numbers, and/or make measurements of quantities and need to write them in a way that indicates how precise the measurements are.

Natural Logs in the Real World

Asked by Lee Hughes, New Lima H.S. in Oklahoma on Sunday Feb 11, 1996:

> I need information on natural logs as it applies to the natural world. This includes such things as plant or population growth or decay such as a bouncing spring. Any information found on the internet or any other resources would be appreciated.

There's a general discussion about some ways in which the number e (on which natural logarithms are based) has relevance in the real world; that discussion is right here on this web site, at `http://www.math.toronto.edu/mathnet/answers/ereal.html`

Any natural phenomenon that grows or decays at a rate proportional to its current value (such as population growth) can be used in place of the compound interest example given in that discussion.

If you want to know some more specific things after reading that, please post another question here and we'd be happy to give additional information.

Use of Neural Networks for Empirical Data

Asked by Domenico Tatone (teacher), Mayfield Secondary School on Friday May 3, 1996:

> I am currently working on a thesis on group dynamics. In my attempt to quantify qualitative research (i.e. interpret responses to interview questions), I am resorting to the development of neural networks. My question relates to the utility of neural networks in empirical studies. Could you direct me to resources that would enable me to implement the proper formation of neural networks, and their respective instruments for testing hypotheses (i.e equations) that emerge from such models. Furthermore, how can fuzzy logic be utilized in conjunction to neural networks.

Here is a reply to your question from a researcher in the Neural Network Research Group here at the University of Toronto:

> It is a little unclear to me what the question is about. But I can certainly recommend an introduction to neural networks. Hertz, Krogh and Palmers "Introduction to the theory of neural computation" is in my opinion by far the best introduction, although I guess a little dated by now. It requires a certain level of mathematical sophistication. Of more recent books there is (physicist) Chris Bishop "Statistical learning in neural networks" and a book by (statistician) Brian Ripley with roughly the same title. Let me know if you have trouble locating any of these and I'll send you proper references.

These general references may not answer your question, but feel free to come back here with more specific questions if you have them.

Calculating Square Roots

Asked by A. Stone and Sonya Baker on January 13, 1997:

> What is the formula for finding square roots?
>
> Thanks

The simplest, but most tedious, way to calculate the decimal digits of a square root is to figure out each digit in turn, as follows.

First of all, find how many digits to the left of the decimal point your square root will have.

A number with one digit to the left of the decimal point will be somewhere from 1 up to, but not including, 10; therefore, its square will be somewhere from 1 up to, but not including, 100, and hence will have either one or two digits to the left of the decimal point.

A number with two digits to the left of the decimal point will be somewhere from 10 up to, but not including, 100; therefore, its square will be somewhere from 100 up to, but not including, 10000, and hence will have either three or four digits to the left of the decimal point.

And so on.

Therefore, if the number you're trying to find the square root of has one or two digits to the left of the decimal point, its square root will have one; if it has three or four, its square root will have two; and so on.

Now, having found how many digits the square root has, just start from the left and figure out each digit in turn.

For example, to find the square root of 1234 to 3 decimal places:

The square root has two digits to the left of the decimal point, so we're looking for a number of the form $AB.CDE\ldots$ (and we want to find A through E).

Now, $10^2 = 100$, $20^2 = 400$, $30^2 = 900$, all too small, but $40^2 = 1600$, too big. This tells us the first digit A is 3.

Next, $31^2 = 961$, $32^2 = 1024$, $33^2 = 1089$, $34^2 = 1156$, $35^2 = 1225$, all too small; $36^2 = 1296$, too big, so the second digit B is 5.

Thirdly, $35.1^2 = 1232.01$, too small, but $35.2^2 = 1239.04$, too big, so the third digit C is 1.

And so on; continuing in this way, you see that $\sqrt{1234} = 35.128\ldots$.

This method is horribly slow and inefficient but that's no problem in an era of calculators. All the necessary ideas are there.

You can apply lots of little shortcuts to this procedure to drastically reduce the number of computations you need to do. Then you come up with a procedure similar to the long-division algorithm. Its complicated to remember, though; I don't remember it offhand. It is seldom taught today because there is little need for it.

To actually calculate a square root to a large number of digits quickly, one uses a different method. For example, saying that $x^2 = a$ is the same as saying $x = (1/2)(x + a/x)$ and you can use this as the basis of an iteration. Start with one value of x, for example, $x = 1$. Then get a second value $y = (1/2)(x + a/x)$. Then get a third value $z = (1/2)(y + a/y)$. And so on. It turns out that these numbers $x, y, z, \ldots$ converge quite quickly to the desired square root, especially if you start with a number that is already close.

Doing this calculation in the above example, for $a = 1234$, and starting with the number $a = 40$ which is obviously just a little bit bigger than the desired square root, you get the numbers $40, 35.425, 35.129578\ldots, 35.128336\ldots, 35.128336\ldots$ and you see that just four calculations are enough to get the square root accurate to six decimal places.

Asked by Sam Alexander (6th grade) on April 30, 1997:

> What is the square root of 2?

The square root of two is the one and only positive real number whose square is 2. That's really the only way to answer the question "what is the square root of 2".

Now, it's possible you may have had a different question in mind: what are the *digits* of the square root of 2?

The digits of an irrational number like the square root of 2 (a number which is not a ratio of two integers) do not ever reach a repeating pattern the way digits of rational numbers do. Therefore, you cannot simply write down *all* the digits of an irrational number, for there are infinitely many of them and they don't repeat. However, you can write down as many as you like, using methods such as that described in the question above. The first few digits of the square root of 2 are 1.41421356 . . ..

Because writing down digits like that doesn't completely and unambiguously identify the number (for instance, I wrote the first 8 digits after the decimal point, but that doesn't tell you what the ninth is), a mathematician would never say that "1.41421356 . . ." is an "answer" to the question "what is the square root of 2". Rather, the correct answer to the question is the first answer I gave above: "the one and only positive real number whose square is 2". That is what the square root of 2 actually *is*. Other responses, such as writing down digits, are useful in understanding the *properties* of $\sqrt{2}$, such as where it fits in to the number line and how it compares with other more familiar numbers, but they are not answers to the question of what it *is*.

Asked by a student at Alfred M. Barbe school on May 1, 1997:

> What is the technical name for the square root symbol (radix)?

It can be referred to as either a "radical sign" or a "surd symbol" (or various other combinations of these, such as "radical symbol").

I believe it was originally derived from a corrupted form of the initial letter of the word "radix", so "radical sign" might be the most appropriate.

Asked by an anonymous poster on October 17, 1997:

> I read the method of calculating square root I wanted to know if there is a way to calculate any degree of root 3,4 and so on what about rational roots 2.5 and so on is there any methods for those? I thank you for your time and if you would be able to send me an e-mail notice when the answer will be posted I will thank you more p.s. your site is the most informative and useful even for university student in math and science department as myself

The methods for computing square roots illustrated above can be generalized to handle any rational exponent. We will first consider the case in which the degree of the root to be taken is a whole number n. Let a be the number of which we wish to take the nth root. As before, first find the greatest power of 10 which is less than the nth root of a. Increase the leading digit of this power of 10 until you reach a number such that if you increase the leading digit one more time, the nth power is greater than a. Move to the next digit and repeat the process.

If a is 7809 and n is 3 then 10 is the greatest power of 10 whose 3rd power (1000) is less than 7809. Increasing the leading digit of 10, we find that 10 cubed is less than 7809 but 20 cubed is greater. This shows our desired root is of the form 1A.BCD. . . . Going to the next digit, we note that 19 cubed is 6859 but 20 cubed is 8000. This shows our desired root is of the form 19.BCD. . . . We continue the process as long as is necessary to obtain the desired number of digits of the root.

If the degree of the root is a fraction, i.e. we are trying to compute $a^{m/n}$ for some integers m and n, then we first take the nth root of a and then raise this number to the mth power. Note that the degree of the root is the reciprocal of the exponent of a: $\sqrt[n/m]{a} = a^{m/n}$.

There is a faster approach which can be used, though the reason why it works is not quite so simple. We first need a rough estimate x of the value of root $\sqrt[n]{a}$. From this we compute

$y = [(n-1)/n]x + a/[nx^{n-1}]$. Now if our first guess was reasonably close to the actual value of the root, then y will be even closer. If we then repeat the process using y for our guess, the new value $z = [(n-1)/n]y + a/[ny^{n-1}]$ which we get will be closer still. The sequence of numbers $x, y, z, \ldots$ obtained in this way converges quite quickly to the desired root, as long as the initial guess was close enough.

How to Draw a Circle on a Computer

Asked by R.S. BHOGAL, teacher, G.R.D. Academy on April 14, 1997:

> I am looking for an algorithm to generate a circle/arc of a given radius R in assembly language. An important factor to note is that I am using 8085 based assembly language for this and it does not have Square Root or Trigonometric functions.

Here are two methods for drawing a circle when there is no access to trigonometric functions or square roots.

One approach is to approximate sine and cosine functions by means of a Taylor series expansion. Usually there is only a need to compute a few terms of the sum. The expansions for sine and cosine are as follows:

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \cdots$$

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + \cdots$$

Here x is in radians (1 radian $= 180/\pi$ degrees). If you truncate the sum at the nth term, the error will be less than the $n+1$st term in the series. (This is true in general for an alternating series in which the absolute values of the terms are strictly decreasing.)

To make the error estimates even better, one can use trigonometric identities such as $\cos(x) = \sin(\pi/2 - x)$ and $\sin(x+\pi) = -\sin(x)$ to ensure that the angles used in these series are between $-\pi/4$ and $\pi/4$ (the closer x is to 0, the fewer the number of terms which are needed to get the desired accuracy). For example, when x is in the range from $-\pi/4$ to $\pi/4$, the expression $x - x^3/6$ approximates $\sin(x)$ with an error of at most 0.0025.

The other method is less mathematical and involves more brute force. Here a routine runs though all of the x and y values of the pixels on screen and sees if the point (x, y) lies on the circle. Actually, since the pixels are discrete, very few will actually lie on the mathematically ideal circle, but instead you can check if they are within a 1/2-pixel distance of the ideal circle by checking if $\sqrt{x^2+y^2}$ is within 0.5 of R. This can be done without the square root function by seeing if x^2+y^2 lies between $(R-0.5)^2$ (which equals $R^2 - R + 0.25$) and $(R+0.5)^2$ (which equals $R^2 + R + 0.25$).

This calculation can be performed using only integer calculations: for each pixel, evaluate x^2+y^2 and see if it is in the range from $R^2 - R + 1$ to $R^2 + R$ inclusive. If so, colour the pixel on the screen, and if not, don't.

Although there are many more computations required for the second method, the calculations are much faster because they are integer calculations. This method would also make it much easier to colour in the interior of the circle or do other similar modifications to the output. The program could be further accelerated if there was some sort of a preliminary routine which decided which points were definitely not worth checking (for instance you would only need to check the points inside the smallest box containing the circle).

Which routine is faster may actually depend on the machine you are using. Some computers with math coprocessors handle decimal computations better than others.

Finding the Focus of a Parabolic Dish

Asked by Nelson Siu, student, Gladstone Secondary on May 4, 1997:

> I'm particular puzzled as to how I can calculate the focus of a parabolic dish. I seem to feel that it is somewhat a reverse of a locus that I've stumbled upon before. I need to find the focus because I'm using this calculation to build a parabolic listening device for my electronics class.

It is easy to find the focus of a parabola (or the three-dimensional version, a "paraboloid") from its equation. If you draw your coordinate axes so that the origin is at the parabola's vertex and the y-axis points in the direction toward which the parabola opens, then its equation will take the form $y = ax^2$ for some constant a, and the focus is the point $(0, 1/(4a^2))$. Similarly for a paraboloid: if the paraboloid opens in the direction of the z-axis, it will have an equation of the form $z = a(x^2 + y^2)$, and the focus is at the point $(0, 0, 1/(4a^2))$.

However, these formulas are not much use in finding the focus of a particular dish you have lying around, since then you would not know the equation of the dish. Rather than trying to make measurements to determine the equation and calculate the focus from those measurements, you'd be better off finding the focus directly by experimentation (e.g., shine several parallel light beams at the dish, each parallel to the main axis of the dish, and observe where they converge). You'd likely get more accurate results that way than trying to determine the equation of the dish by measurements.

The Decomposition of a Drug in the Human Body

Asked by Irene Funk, student, Newark High School on June 16, 1997:

> How can I determine the rate at which a drug is broken down by the human body?

Unfortunately your question concerns biology more than mathematics, and we simply don't have the resources to answer non-mathematical questions.

You would need to know the biological principles specific to the situation you have in mind, principles which would relate the rate of decomposition to other factors. These principles would then be expressible as mathematical formulas involving the quantities of the drug, and the problem would enter the realm of mathematics at that point.

If you have such formulas and have a mathematical question about how to use the formulas to determine other things (such as how to get from a formula describing the rate of decomposition in terms of the current concentration into a formula describing the concentration as a function of time), please post another question here and we'd be glad to address it.

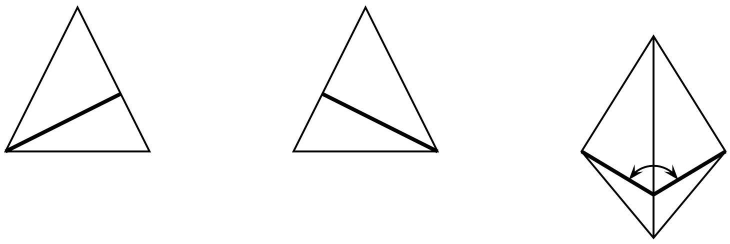
Calculating Angles In A Pyramid

Asked by Clive Champion on July 31, 1997:

> I am an artist and have a project where I have to construct a 4 sided pyramid out of glass for a garden fountain. Each of the four triangles have a base of 30.5 inches and a height of 46 inches. I need to know what degree of miter to put on the edges of the uprights of the triangles so that it will all fit together nicely. Thanks for any help you can give me. I have come up with 54.125 degrees on each of the upright edges. I am sure there must be a formula but I cannot find it. Thanks again.

Let's denote the measure of the angle of the miter by A. The angle of the miter needs to be half the angle formed by two faces of the pyramid, so the angle between these faces is $2A$. If we find this angle, we can divide by 2 to get A.

The way to find the angle between these two adjacent triangular faces is to think of the common edge of these triangles as their bases and to construct two line segments, perpendicular to this edge, corresponding to their altitudes. The angle that these line segments make when joined together is the same as the angle between the faces, which is $2A$. This is illustrated below (the first picture shows the triangular faces separate, the second shows them assembled into the pyramid).



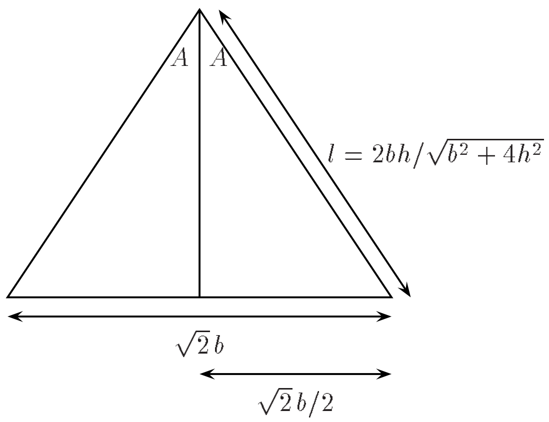
Now we need to compute the angle at which these two line segments meet. To do this we need two things: (a) their length (they are of equal length), and (b) the length of the third side of the triangle formed by these two line segments.

Let's find length (b) first. The third side (running from the end of the second line segment back to the start of the first) is a diagonal of the base of the pyramid. Since the base of the pyramid is a square, whose sides are the bases of the triangles, the length of this diagonal is $\sqrt{2}b$ where b denotes the base length of each triangle.

Length (a) be computed by finding the area of a face of the pyramid in two ways. One way is to use the standard base and height measurements (thinking of the bottom edge as the base). The area is $1/2bh$, where h is the height and b the length of the base.

The other way is to think of one of the slanted edges of the triangle as the "base", with our unknown line segment length being the "height". The length of the edge can be computed, using the Pythagorean theorem, to be $\sqrt{h^2 + b^2/4}$. If we let l denote the length of our unknown side, we can see that the area is $1/2l\sqrt{h^2 + b^2/4}$. Since we also know that the area is $1/2bh$, we can equate these two values and solve for l, which gives $l = 2bh/\sqrt{b^2 + 4h^2}$.

We now know that our two line segments, which meet at an angle of measure $2A$, have length $l = 2bh/\sqrt{b^2 + 4h^2}$ and form two sides of a triangle whose third length is $\sqrt{2}b$. As illustrated below, that means the angle A is part of a right-angled triangle whose hypotenuse has length $l = 2bh/\sqrt{b^2 + 4h^2}$ and whose opposite side has length $\sqrt{2}b/2$.



Thus the angle A is given by $\arcsin(\sqrt{2}\,b/(2l)) = \arcsin(\sqrt{8h^2+2b^2}/4h)$. Plugging in our values of $h = 46$ and $b = 30.5$ we find that A is approximately 48.155 degrees.

Applications of Polynomial Factorization

Asked by Blanche Reyna on October 4, 1997:

> I need help on how to justify factoring of polynomials. However, I would like an explanation that can be understood by students in pre-Algebra? Can anyone help?

I'm not entirely sure what you mean when you want to "justify" factoring of polynomials. Do you mean a mathematical argument that shows why techniques used to factor polynomials are correct and give the right answer, or are you asking what use it is to know how to factor polynomials?

I will assume you are asking the latter, looking for reasons why it is useful to know how to factor polynomials.

The most fundamental reason is to be able to solve polynomial equations. If you have an equation saying that a product of several factors must equal zero, the solution is that one or more of the factors must be zero.

For example, suppose you throw a ball into the air and want to find when it hits the ground. Laws of physics tell you that the height of the ball t seconds after you threw it is a quadratic polynomial in t. Depending on how fast and at what angle you threw it, you might, for instance, discover from the laws of physics that the height of the ball is $-t^2+4t+5$. To find when it hits the ground, you want to find what value of t gives you $-t^2+4t+5=0$. A tough problem as it stands (until you learn about the quadratic formula). But if you're able to factor the left-hand side, you see that the equation is saying that $(-t+5)(t+1)=0$ and the only time this product is zero is when one of the factors is zero; in other words, when $t=5$ or $t=-1$. Ignoring the negative solution (because the ball didn't start moving until after you threw it), you see that the ball hits the ground five seconds after you threw it.

You could explain to your students that they will learn a formula for solving any quadratic equation like this, but that factoring is the way such a formula comes about.

Then you could also mention that knowing how a number factors tells you many useful things about it. For example, why is it that if you take any number and take the sum of that number, its cube, and three times its square, you always end up with a multiple of 6? The answer is because the polynomial x^3+3x^2+2x factors as $x(x+1)(x+2)$, and among any three consecutive numbers at least one must be even and at least one must be a multiple of 3, so the product is a multiple of 6.

Interpreting the "Expected Number"

Asked by Steve Yuen on Monday May 27, 1996:

> We are having a problem with the solution to the following problem.
>
> QUESTION:
>
> The DJ of the local radio show conducts a contest in which the listeners phone in the answer to the contest question. If the DJ estimates that the probability of a caller giving the correct answer is 0.28,
>
> a) What is the probability that the third caller is the winner?
> b) What is the expected number of calls BEFORE a correct answer is received?
>
> Part A
>
> Calculation: 0.72 X 0.72 X 0.28 =0.145152
>
> Therefore, the probability that the third caller is the winner is 0.145152
>
> Part B
>
> The formula of the expected waiting time is $E(X) = q/p$ (Geometric Distribution). p is the probability of success on each trail. $q = 1 - p$ is the probability of failure on each trail. $X = 0, 1, 2, \ldots$
>
> Calculation: By the formula $E(X) = q/p$
>
> The expected waiting time = 0.72/0.28 is about 2.571428571
>
> Here is where the interpretation problem comes
>
> Here are two opposing answers:
>
> Since the expected waiting time is greater than 2.5 then we expect the third caller to get the correct answer. Therefore, the expected number of the calls before a correct answer = 2
>
> Solution # 2
>
> Since the expected waiting time (average number of failures before success) is closer to 3 therefore more often than not the fourth caller will be the first successful one. In this case we have waited through 3 calls.
>
> Which (if any) do you think is the correct answer? (2 or 3)
>
> Thank-you for your help

I think the issue here is the meaning of "expected number of calls". The correct answer to the question that was asked is the number 2.57... that you calculated; that's the expected number of calls.

It's important to realize that this "expected number" is *not* an integer (so it's wrong to round it to either 2 or 3); it is not the actual outcome of any contest. Rather, it means that if you run the contest a large number of times (say around 100 times), you'd expect the total number of calls (failures before success) for all these contests to be around 257; if you ran it 1000 times, you'd expect the number to be around 2571; and so on.

It's like saying that the average number of children per family is 2.5; that doesn't mean that "the average family has 2.5 children". It doesn't mean that a typical family you meet on the street will have 2.5 children; that can't happen, since families have whole numbers of children. Instead, it refers to the average over all families.

Both Solution #1 and Solution #2 are attempting to interpret "expected value" as "what do I expect will happen in one particular contest?", and the expected value doesn't say what you might think it says about that. Here are three examples to illustrate this:

1. Suppose that half of the time there are 2 calls and the other half of the time there are 3 calls. The expected number of calls is 2.5.
2. Suppose that three quarters of the time there are 2 calls and one quarter of the time there are 4 calls. The expected number of calls is $(3/4)(2) + (1/4)(4) = 2.5$.
3. Suppose that one quarter of the time there is 1 call and three quarters of the time there are 3 calls. The expected number of calls is $(1/4)(1) + (3/4)(3) = 2.5$.

In each case there is the same expected number of calls. But what's most likely to happen is very different for each situation: in situation 2, it's most likely that there'll be only 2 calls (this happens 75% of the time), while in situation 3, it's most likely there'll be 3 calls (this happens 75% of the time). So the expected value does not tell you what's most likely to happen.

To put it another way: suppose you got 1 call 99% of the time, and 99901 calls 1% of the time. The expected number of calls is 1000, even though the most likely outcome is only 1 call, with 99% probability! A large number of calls occurring with low probability can contribute just as much to the expected value as a few number of calls occurring with high probability.

If you really want to think about what's most likely to happen in the DJ example, you need to think about the individual probabilities and *not* the expected number:

> Probability of 0 calls before first successfull call: 0.28
> Probability of 1 call before first successfull call: 0.2016
> Probability of 2 calls before first successfull call: 0.1451...
> Probability of 3 calls before first successfull call: 0.1045...
> Probability of 4 calls before first successfull call: 0.0752...
> and so on.

So the most likely outcome is that the first call is successful, the next most likely outcome is that there's 1 unsuccessful call, and so on. The expected number of unsuccessful calls (the average number if you run the contest many times) is 2.57....

What are Antiatoms Made Of?

Asked by Keith Cannon, student, Union Central on March 13, 1997:

> What would an "antiatom," made up of the antiparticles to the constituents of normal atom, consist of? What might happen if antimatter, made of such antiatoms, came in contact with our normal world of matter?

Before explaining what antiatoms are, it would benefit some of our readers to explain what particles and antiparticles are. As most of us know, most matter that we can see is made up of atoms which are in turn composed of various combinations of three particles: electrons, protons, and neutrons. These three particles, together with a wide variety of other more obscure particles, are the basis for all matter.

It has been found that for each of these particles, there is a dual "antiparticle." The dual of the electron is called the positron, while the duals of the proton and the neutron are simply called antiprotons and antineutrons. Antiparticles behave very much the same as ordinary particles. They have the same mass as their particle counterparts, but the charge of an antiparticle is opposite the charge of the corresponding particle.

Just as protons, electrons, and neutrons often cluster together to form atoms, antiprotons, positrons, and antineutrons can combine to form antiatoms. The properties of antimatter are very similar to that of normal matter. While only very basic antiatoms like antihydrogen have been made in laboratories, some scientists believe that there may be entire galaxies composed almost entirely of antimatter.

Perhaps the most interesting thing about antimatter is what it does when it comes in contact with normal matter. When a particle collides with its antiparticle counterpart, the two annihilate each other and give off a burst of energy. The amount of energy given off is given by Einstein's famous formula $E = mc^2$ where E is energy in Joules, m is the mass in kilograms and $c = 3 \times 10^8$ m/sec is the speed of light. Similarly, when atoms and antiatoms collide, they also destroy one another and release energy. To give an idea of just how much energy is produced, consider a paper clip (about 1 g in mass) combining with an antipaperclip. The energy produced would be $E = (2 \times 10^{-3})(9 \times 10^{16}) = 1.8 \times 10^{14}$ Joules. For the sake of comparison, it takes about 3×10^9 Joules to power a 100 Watt light bulb for a year.

Followup Question by Brent Potteiger on March 30, 1997:

> I understand the concept of $E = mc^2$, and after reading your article about antiatoms (and figuring out that two grams of matter could create enough energy to power approximately 60,000 100 kilowatt light bulbs for a year) I don't understand why we don't harness the power to convert matter completely into energy. Wouldn't that save all of our energy needs?

At the present we do not have an efficient way of making antimatter. To harness the energy in matter we would need to take one of two approaches.

One would be to generate antimatter. The only known methods of accomplishing this involve starting with energy and then converting it into antimatter. The methods for doing so are very inefficient and are only of use (so far) in that they allow us to study the behavior of antimatter. The energy resulting from the antimatter created would be far to little to make up for the massive amounts of energy it took to generate the antimatter. Also, at the present, the amount of antimatter that can be generated by even a very large collider is on the atomic level. Even if we could efficiently generate antimatter, we would need to be able to do it on a much larger scale.

Another approach would be to discover a source of naturally occurring antimatter. While natural antimatter likely does exist, it is likely (and hopefully) far away. Because it reacts so readily with matter, it is typically not found anywhere near matter.

Another problem is the means by which one might contain the antimatter and then capture the energy once it reacts with matter. Containing antimatter is no small feat since it must be stored in something which is not made of matter. Currently electric and magnetic fields are used, though it is not clear how one would use this sort of containment device on any more than a few atoms or particles.

The energy produced in antimatter/matter reactions is in the form of gamma rays. This is similar to the case of fission reactor, where water is heater and used to drive a turbine. The problem is that, even though there is no inherent radioactivity in the matter or antimatter, anything which comes into contact with the reaction will eventually become radioactive. There is no waste, though there is still the problem of what to do with an old power plant which has itself become radioactive over the years.

As a closing remark, all the basic theory is in place for extracting energy from nuclear fusion. A fusion reaction is much easier to implement than a matter-antimatter reaction but is still a long way from perfection. Some experimental fusion reactors have been built, but currently they are, at best, breaking even (they require as much energy to run as they produce). The problem is not with fusion as an energy source, but rather with the practical aspects of how to get the reaction going and then efficiently harness energy from it.

Followup question by an anonymous student on September 28, 1997:

> So, why do matter and antimatter release energy when combined?

One of the great breakthroughs of this century was the discovery that mass was itself a type of energy. This energy can be converted into other, more conventional sorts of energy only in special cases though. One of these instances is a matter/antimatter reaction (other examples are nuclear fusion and fission). When matter and antimatter destroy one another, they must release the energy that was stored in their mass (energy is always conserved). This energy is typically in the form of a very high energy light particle called a gamma ray. The amount of energy which is stored in the mass of a particle is given by Einstein's famous formula $E = mc^2$ (E is energy in Joules, m is mass in kilograms, and c is the speed of light in a vacuum in meters per second).

The Four Fours Problem

Asked by Hazen Simson, Grade 12, RCS-Netherwood on October 26, 1996:

> I would like to know if you know the answer to the four number problem that mathematicians have done in previous years and all mathematicians know of. It is something I have been trying for several months and without being able to get above 10 I am now searching for an answer.

I'm not 100% sure what you're asking about, but I assume (from your remarks about not being able to get above 10) that you're referring to the problem about representing integers using four 4's and operations such as addition, multiplication, etc. (If I'm wrong and you're referring to something else, such as the four-colour theorem, please let us know).

It's not really a problem "that mathematicians have done in previous years and all mathematicians know of" because it isn't really a mathematical problem. It depends on what notational convention one uses, rather than any kind of mathematical truths.

In its most basic form, the puzzle is to combine four copies of the number 4, through the basic operations of negation, addition, subtraction, multiplication, division, and exponentiation, to come up with different integers.

In this form, you can represent all the integers from 1 to 9, and some others, but you cannot represent 10. (You can prove that you can't by having a computer generate all the possible combinations and list the results).

However, there are several different ways to change the rules of the puzzle.

One way is to allow the square root symbol, so that you can take square roots without using up any additional fours (instead of having to raise something to the power of $4/(4+4)$ which is how you'd have to do it under the original rules).

With this, you *can* represent 10 (as $4+4+4/\sqrt{4}$). However, I don't believe you can represent 11 in this manner. (Mind you, you can't prove that 11 is impossible in the same way you can prove 10 is impossible under the original rules, because there is no limit on the number of times the square root symbol can appear in your formula, so there are an infinite number of possible expressions instead of a finite number and you can't just have a computer check them all. A proof would have to involve more advanced ideas from abstract algebra).

Another way to change the rules is to allow, not just combinations of the number 4, but combinations of numbers that can be written in decimal notation using the digit 4 (for example, 44 or .4). With these rules, you can represent $10 = 4/.4+4-4$, $11 = 44/4+4-4$ (or $11 = 4/.4+4/4$), and $12 = (44+4)/4$. However, you cannot represent 13.

Another way to change the rules is to allow all of the above. This is probably the form of the puzzle you are working on. You still can't generate all integers, but you can go quite a bit past 13. The previous examples should help you figure out how.

Finally, you could change the rules to allow functions such as logarithms to occur in your expression. If you do that, then *every* integer can be expressed. For example:

$$\begin{aligned} 1 &= -\log_{4/\sqrt{4}} \log_4 \sqrt{4} \\ 2 &= -\log_{4/\sqrt{4}} \log_4 \sqrt{\sqrt{4}} \\ 3 &= -\log_{4/\sqrt{4}} \log_4 \sqrt{\sqrt{\sqrt{4}}} \end{aligned}$$

and so on. (See if you can prove this!)

In summary, this puzzle depends entirely on what rules you choose. It has more to do with manipulating the symbolism in clever ways than on any mathematical truths, and if mathematical notation had evolved differently, the outcome of the puzzle would be quite different too.

Generalizing the Towers of Hanoi Problem

Asked by David Watts on November 10, 1997:

> I am currently studying the towers of hanoi problem. I know that if you increase the number disks you increase the minimum number of moves (using $2^n - 1$ to find out how many), but what would happen if you:
>
> 1: Increased the number of pegs? (e.g., 4 pegs, so you have to move all the discs from Peg 1 to Peg 4)
>
> or
>
> 2: Only moved one peg at a time? (ie $1 \to 3$ would not be allowed, but $1 \to 2 \to 3$ would be)
>
> or
>
> 3: Or even a combination of the above? (e.g., $1 \to 2 \to 3 \to 4$)
>
> Can anyone come up with any ideas for those? Any predictions or answers would be most gratefully accepted.
>
> Thanks in Advance
>
> David Watts

Let's start with your second question, where the pegs are lined up in a row and you can only move between adjacent pegs. This can be solved by an inductive argument similar to the usual case.

For example, let $f(n)$ be the minimum number of moves required to move a pile of n disks from one of the outer pegs to the other outer peg (such as from peg 1 to peg 3, or from peg 3 to peg 1). To accomplish a move of n disks from peg 1 to peg 3, the bottom disk will eventually have to move from peg 1 to peg 3. It can't do so in one move, so it must first move from peg 1 to peg 2 and then move from peg 2 to peg 3. Thus the bottom disk must make at least 2 moves.

Before the bottom disk can make its first move (from peg 1 to peg 2), the top $n-1$ disks must all be on peg 3. It will take $f(n-1)$ moves to get them there. Next, before the bottom disk can move from peg 2 to peg 3, the top $n-1$ disks must all be moved from peg 3 back to peg 1; this takes another $f(n-1)$ moves. Finally, the top $n-1$ disks must be moved to peg 3 again, requiring a third sequence of $f(n-1)$ moves. Thus the top $n-1$ disks must make a total of at least $3f(n-1)$ moves.

The total number of moves required is therefore $f(n) = 3f(n-1)+2$. This recursion formula can be solved in the same manner described in the discussion of the standard problem, yielding the solution $f(n) = 3^n - 1$.

Next, you could ask about moving the pile not from one outer peg to the other, but from an outer peg to an inner peg: say from peg 1 to peg 2. Let $g(n)$ denote the number of turns required to do this for a pile of n disks. (By symmetry this is the same as the number of turns required to move the pile from peg 3 to peg 2, but it is not obvious whether or not this is the same as the number of turns required to move the pile from peg 2 to peg 1 or from peg 2 to peg 3.)

At some time the bottom disk must move from peg 1 to peg 2. Before this can happen, the top $n-1$ disks must have been moved from peg 1 to peg 3; this takes $3^{n-1}-1$ turns. Then the top $n-1$ disks must be moved from peg 3 to peg 2; this takes $g(n-1)$ turns. Hence we have

$$\begin{aligned}
g(n) &= 3^{n-1} + g(n-1) \\
&= 3^{n-1} + 3^{n-2} + g(n-2) \\
&= \ldots \\
&= 3^{n-1} + 3^{n-2} + \cdots + 3^1 + g(1) \\
&= 3^{n-1} + 3^{n-2} + \cdots + 3^1 + 1 \\
&= (3^n - 1)/(3-1) \\
&= (3^n - 1)/2
\end{aligned}$$

(using the formula for the sum of terms in a geometric progression).

Finally, if you do a similar sort of analysis, you'll discover that this does also happen to equal the number of turns required to move the pile from the middle peg to one of the outer ones.

Now on to your first question. The general case is unsolved. It is not known what the minimum number of moves is as a function of the number of disks and number of pegs.

Some parts of the problem are known; for example, if the number of pegs is greater than or equal to the number of disks, it is easy to see that the minimum number of moves is $2n-1$.

For other configurations there are various unproven conjectures. For example, in the case of 4 pegs, it is conjectured (but not known for sure) that the optimum strategy for moving n disks is to first move some of the topmost disks (say the top k disks) to one of the spare pegs; this takes $f(k)$ moves. Then there are 3 pegs available; use the 3-peg strategy to move the remaining $n-k$ pegs to their destination (which takes $2^{n-k}-1$ moves). Finally, move the top k disks into their final destination, which takes $f(k)$ moves.

Therefore, for any choice of $k < n$, it is possible to move n disks in $2f(k) + 2^{n-k} - 1$ moves. It is conjectured that, if you take the value of k that requires the least number of moves, this is an optimal strategy, so

$$f(n) = \min_{k<n} 2f(k) + 2^{n-k} - 1.$$

It is not known for sure that this is always optimal, though.

If you calculate values of this function f, you obtain

n	1		2		3		4		5		6		7		8		9		10		11
f(n)	1		3		5		9		13		17		25		33		41		49		65
d		2		2		4		4		4		8		8		8		8		16	

Of particular interest are the differences between consecutive $f(n)$ terms (shown in the last row of the table): two 2's, then three 4's, then four 8's, then five 16's, etc.

This pattern allows you to find a formula for $f(n)$. To find this formula, it may be easiest to think first of values just before the gap jumps up: at $n = 1$, at $n = 1+2 = 3$, at $n = 1+2+3 = 6$, at $n = 1 + 2 + 3 + 4 = 10$, etc. If we let $g(N)$ denote $f(1 + 2 + \ldots + N)$, then $g(N)$ equals $g(N-1)$ plus N gaps of size 2^{N-1} each, so $g(N) = g(N-1) + N2^{N-1}$.

There are various techniques to solve this recursion formula and get an explicit formula for $g(N)$. One way, not the most elegant but useful if you aren't familiar with standard techniques, is to write out

$$\begin{aligned} g(N) &= N2^{N-1} + g(N-1) \\ &= N2^{N-1} + (N-1)2^{N-2} + g(N-2) \\ &= \ldots \\ &= N2^{N-1} + (N-1)2^{N-2} + \cdots + 2\ 2^1 + g(1) \\ &= N2^{N-1} + (N-1)2^{N-2} + \cdots + 2\ 2^1 + 1 \end{aligned}$$

and then split this up as

$$\begin{aligned} g(N) = \; & 2^{N-1} + 2^{N-2} + \cdots + 2^2 + 2^1 + 2^0 \\ & +2^{N-1} + 2^{N-2} + \cdots + 2^2 + 2^1 \\ & +2^{N-1} + 2^{N-2} + \cdots + 2^2 \\ & + \cdots \\ & +2^{N-1} + 2^{N-2} \\ & +2^{N-1} \end{aligned}$$

(there are N lines here, with 2^{N-1} appearing in each of them for a total of $N2^{N-1}$, with 2^{N-2} appearing in all but one of them for a total of $(N-1)2^{N-2}$, and so on).

Now rewrite the above as

$$\begin{aligned} g(N) = \; & 1 + 2 + \cdots + 2^{N-1} \\ & +2(1 + 2 + \cdots + 2^{N-2}) \\ & +2^2(1 + 2 + \cdots +^{N-3}) \\ & + \cdots \end{aligned}$$

$$+2^{N-2}(1+2)$$
$$+2^{N-1}(1)$$

If you know the formula for the sum of a geometric series, you know that $1+2+2^2+\cdots+2^{k-1} = 2^k - 1$, so we have

$$\begin{aligned}
g(N) &= 2^N - 1 \\
&\quad + 2(2^{N-1} - 1) \\
&\quad + 2^2(2^{N-2} - 1) \\
&\quad + \cdots \\
&\quad + 2^{N-1}(2^1 - 1) \\
&= (2^N - 1) + (2^N - 2) + (2^N - 2^2) + \cdots + (2^N - 2^{N-1}) \\
&= N2^N - (1 + 2 + \cdots + 2^{N-1}) \\
&= N2^N - (2^N - 1) \\
&= (N-1)2^N + 1.
\end{aligned}$$

This proves that $g(N) = (N-1)2^N + 1$. (There are much shorter ways to get this, but this is probably the most elementary).

Now, to get a formula for $f(n)$, let N be the largest number for which $1 + 2 + \cdots + N$ is less than or equal to N. Write $n = (1 + 2 + \cdots + N) + m$; then $f(n)$ differs from $f(1 + 2 + \cdots + N)$ by m gaps of size 2^N, so

$$\begin{aligned}
f(n) &= f(1 + 2 + ... + N) + m2^N \\
&= g(N) + m2^N \\
&= (N-1)2^N + 1 + m2^N \\
&= (N + m - 1)2^N + 1.
\end{aligned}$$

Now the only task left is to express N and m as functions of n. To do this, use the formula $1 + 2 + \cdots + N = N(N+1)/2$; N is the largest integer such that $N(N+1)/2 \leq n$. In other words, it is the largest integer less than or equal to the positive root r of the quadratic equation $r(r+1)/2 = n$. Use the quadratic formula to express r as a function of n; then $N = [r]$ (greatest integer less than or equal to r), and $m = n - (1 + 2 + \cdots + N) = n - N(N+1)/2$.

Plugging all that into the formula for $f(n)$ gives

$$f(n) = ([r] - (n - [r]([r]+1)/2) - 1)2^{[r]} + 1$$

and setting r to be what you find by the quadratic formula gives you $f(n)$ as a function of n.

Now, this is under the assumption that the observed pattern of gaps really does always continue to hold true. To prove that it does, what you need is to go back to the original definition

$$f(n) = \min_{k<n} 2f(k) + 2^{n-k} - 1$$

and prove by induction that the formula for f(n) found above satisfies this equation. It's not an easy task.

Finally, remember that all this is only for the case of 4 pegs, and is only for the strategy described above (moving a chunk of disks first, then moving the remainder using the 3-peg strategy, then moving the chunk back). This strategy is conjectured, but not known, to be optimal. So the above formula for $f(n)$, even after you prove that it matches our definition, is still only conjectured but not known to be the smallest number of moves required in the 4-peg case.

For other numbers of pegs you can perform similar analyses but things become fare more complicated far more quickly. Again, there are conjectures, but the minimal number of moves is not known (except for small numbers of disks where an exhaustive computer search is possible).

As to your question 3 (the 4-peg case where there is the additional constraint of only being able to move to adjacent pegs), I imagine that too would be a very complex situation for which the exact solution is not known. I have not given it much thought, though.

Asked by a student at Oak Park High School on January 17, 1998:

> About the Towers of Hanoi ...
>
> 1. It has three pegs(let's name them S, D and A)
> 2. All the normal rules are the same except...
>
> No moves are allowed from peg S onto peg D, directly (although the moves from D to S are allowed.)
>
> My question is that what's the minimal number of moves can make??
>
> Would you please give me a hand about this problem?
>
> Thank you very much

This problem is similar to case 2 (the easier case) of the generalization asked in the previous question.

Let $f(n)$ stand for the minimum number of turns it takes to move a pile of n disks from peg S onto peg D. Let $g(n)$ stand for the minimum number of turns it takes to move a pile of n disks from peg D onto peg S.

To accomplish a move of n disks from peg S to peg D, the bottom peg will eventually have to move from peg S to peg D. It cannot do so directly, so it must first move to peg A then to peg D. Therefore, the bottom disk must make at least 2 moves.

Before the bottom disk can move to peg A, the top $n-1$ disks must all be moved from peg S to peg D. The minimum number of turns required for this is $f(n-1)$.

Before the bottom disk can move from peg A to peg D, the top $n-1$ disks must all be moved from peg D back to peg S. The minimum number of turns required for this is $g(n-1)$.

Finally, the top $n-1$ disks must be moved back to peg D, requiring $f(n-1)$ turns.

Therefore, the total number of moves required to move the pile of n disks is $2+2f(n-1)+g(n-1)$, so we have the recursion formula $f(n) = 2 + 2f(n-1) + g(n-1)$.

If you do a similar analysis for moving a pile of n disks from peg D to peg S, you get the recursion formula $g(n) = 2 + 2g(n-1) + f(n-1)$.

One way to solve this pair of linked formulas is to let $F(n) = f(n) + g(n) = 2 + 2f(n-1) + g(n-1) + 2 + 2g(n-1) + f(n-1) = 4 + 3[f(n-1) + g(n-1)] = 4 + 3F(n-1)$. This you solve using the same technique as described above.

Next, let $G(n) = f(n) - g(n) = 2 + 2f(n-1) + g(n-1) - 2 - 2g(n-1) - f(n-1) = f(n-1) - g(n-1) = G(n-1)$. Since $G(n) = G(n-1)$ for all n, it follows that $G(n) = G(1) = f(1) - g(1) = 2 - 1 = 1$.

Finally, you can solve for $f(n)$ and $g(n)$ using the fact that $f(n) - g(n) = 1$ to write $g(n) = f(n) - 1$, and plug that into $f(n) + g(n) = F(n)$ to get $f(n) = (F(n) + 1)/2$. I'll leave it to you to finish up the argument and solve for $F(n)$.

Is Deductive Geometry Worth Salvaging in the High-School Curriculum?

From Alexandru Pintilie (teacher), Bayview Glen School:

> The study of deductive geometry in high schools is almost inexistent. Time and the students' level of abstraction make it almost impossible to teach. I have also met math teachers who "hate proofs". Is deductive geometry worth salvaging or do we move towards other areas: optimization, probabilities etc...?

From Philip Spencer, University of Toronto:

> I have seen many students who cannot follow the course of a logical argument, and who cannot make the connection between a series of symbol manipulations by which one arrives at "an answer" and the progression of ideas that turns those manipulations into a convincing proof.
>
> My own personal opinion is that it would be highly desirable to restore to the curriculum either deductive geometry, or some similar subject, in which students can see the relationships between tightly interconnected ideas, learn the basics of axiom systems and proofs, and encounter rigorous reasoning.
>
> However, perhaps the demise of the subject is an indication that improvements need to be made to it rather than restoring it wholesale to its original state. Certainly Euclid's *Elements*, while an admirable feat of logic and reasoning, leaves much to be desired pedagogically. There's a vagueness of definition and lack of motivation, leaving students with the feeling that calling something an axiom is simply a crutch when you can't think of a good way to define or justify it, that the laws of geometry are a dry collection of incomprehensible theorems without any relevance, and many other reactions which have contributed to the dislike of the subject by students and teachers alike.
>
> Perhaps our challenge is to rework the material in a way which is less susceptible to these misconceptions; that highlights geometric theorems as things known to be true with greater certainty than simply because they've been true in all the examples we've happened to observe, being instead necessary logical consequences of other things we know to be true; that establishes this way of thinking as useful and necessary in many other unrelated fields as well, with geometry merely a particularly cogent and complete instance of it; that illustrates the process of discovery of new truths through reasoning; and that is more appealing to student and teacher alike.
>
> Does anybody have any thoughts on how this can be done?

From Edwin Sherman on December 18, 1996:

> I am really sorry to hear that Question even being asked but it is good that it was because it shows the sorry state of the teaching. I do agree that it is difficult to teach this subject and it always will be as long as it is left unmentioned until you get to Geometry in High School. Deductive reasoning is a BASIC part of all Math, (and all life) not just Geometry. For the last 50 years, and probably earlier, Math has been taught by memorizing tables whether it be, addition, substraction, multiplication or division. What has happened is students know $2 + 2 = 4$ but don't know why.
>
> Deductive reasoning must be started right at the beginning in the first grade. It starts with What is zero, and why does one plus one equal two. As long as math

is taught with memorization only, we will have lots of people that will be able to obtain the answer to very complex problems but won't have the slightest idea why it is so. The grade and high schools believe that deductive reasoning is used only in higher math so they leave it to the colleges for teaching. Meanwhile, the colleges are expecting all entering students to already have a sound foundation in it. It goes back to the old adage, you can feed a hungry man or teach him how to fish. Feeding the students formulas, no matter how complicated, is still the feeding end where the deductive reasoning is the teaching of how to fish. This doesn't mean you start with the teaching of geometry although that would help but it can be taught also with simple basic math.

For example, if you have 24 apples in a box, 10 boxes on a pallet and three pallets of apples, the deductive reasoning part is teaching you can't multiply the apples times the pallets without including the boxes. That is where you teach students how to think. That is deductive reasoning.

Until this is recognized we will continue with the sorry state of math general knowledge that we have. Trouble is, we have traveled this path so long that even many of our grade school teachers cannot now answer a student's question of "Why is two and two, four?" Too often it is just answered, "Just because it is."

Deductive reasoning is used in every part of life. It is used in deciding what to buy, what to wear, even what to eat. It is basically simply how to think and knowing how we come to that conclusion and why. That should be the main purpose of education, not memorizing dates, names or even what happen but why it happened. What has happened is, we now have a generation that know how to use a can opener but are perplexed by not knowing why it won't work on a 55 gal drum.

No, deductive reasoning should not be dropped. Instead it should be started in the first grade and it wouldn't hurt if it was a separate course. Sorry to rattle on but this is to important to not speak out.

Followup comment by Colin Dawson, grade 9 on December 30, 1997:

I am currently taking geometry. I think that deductive reasoning doesn't need to be further developed in this course because I feel that deductive resoning is a skill that can be naturally attained through your everyday life. Geometric Proofs only confuse students which furthers their personal beliefs of the mundane nature of geometry.

Anyone Have Interesting Graphing Calculator Problems?

From David Lederman, "David Lederman's Math Club":

If you are a teacher of calculus and are using graphing calculators in teaching calculus, I would like to hear from you. Also, if you have any interesting ways of presenting any topic(s) in calculus or have interesting graphing-calculator problems, please communicate with me at lederman@nwdc.com.

Editor's Note: If you have such ideas, please also post them here so that everybody can benefit from them and contribute to the discussion.

From Alex Pintilie (teacher), Bayview Glen School:

I think that it would be quite neat to illustrate Newton's Method of finding the roots of an "ugly" equation using graphing calculators. Take a point on the curve, draw the tangent, cut the x-axis. Zoom in, take the tangent at the new point and so on. Students will (hopefully) understand better what hides behind the numerical iterations. This procedure can also illustrate why chosing some "bad" starting point, the method does not converge.

From bassman@recorder.ca, Grenville Christian College, Brockville, ON:

> As an OAC Algebra student,I have occasional use for a graphing calculator. I own a TI-85 by Texas Instruments and I am wondering if it is possible to get it to reduce function equations to a form in which the y variable is expressed in terms of x since that is the only format that it accepts to graph a function. Apart from doing it by hand, I haven't been able to make it do this. Any ideas would really be appreciated.

Reply by Philip Spencer, University of Toronto:

> There is no general procedure to solve an equation for one of the variables. There are two issues involved: first, does such a function even exist, and secondly, if it does, can it be expressed in terms of "familiar" functions like polynomials, exponentials, trig functions, and so on.
>
> First of all, an equation in x and y may not always define y as a function of x. For instance, the equation $x - y^2 = 0$ does not define y as a function of x: for $x = 1$, there are *two* y-values (1 and -1) which satisfy the equation, so y is not uniquely determined as a function of x.
>
> There is a way to tell in most cases whether or not an equation defines y as a function of x, but it involves notions from multi-variable calculus so I won't get into it in detail here, unless someone wants me to.
>
> However, even if you know that an equation does define y as a function of x, there is no guarantee that this function can be expressed in terms of familiar functions like polynomials, exponentials, trig functions, and so on. The best you can do is try various "by hand" techniques of solving the equation for y, and see if any of them work.
>
> Because this involves some intelligent guesswork and choosing of appropriate manipulations, it is beyond the scope of a graphing calculator. A more sophisticated computer algebra program like Mathematica or Maple will be able to do it in many cases, but even then it's not guaranteed to work.

How Much Does School Size affect Performance in Math Contests?

From Alex Pintilie, Bayview Glen School:

> This is surely controversial enough to be discussed. Recently, we received the results from the Waterloo math competitions. Of course, participation is what matters but we would have liked to see our school's name a bit higher in the rankings. Being a relatively small school, we can use the excuse that it easier to find 3 math wizards in a school with 200 grade 9 students than in a school with 30 grade 9 students.
>
> The topic that I propose for discussion is "How much does the size of a school influence its rank in (say) the Pascal contest?"
>
> I did a little research on my own by "playing" with order statistics in normally generated samples and I found out that increasing the size of our school 3 times

> would have improved our Cayley score from 309 to about 340. Increasing the size 7 times would have brought it only up to 350. It seems to matter if the school is really small but after a certain size it does not seem to matter as much as I would have expected it.
>
> Has anybody given this topic a thought?

From David Dymov, May 28, 1997:

> I think it doesn't matter how many students there are. What is more important is the quality of the teaching staff.

From Dorian Smith, Arizona State University, November 14, 1997:

> Are there any competitions above the high school level?

(From Philip Spencer, University of Toronto) Yes, there is the William Lowell Putnam Mathematics Competition. You can find more information at http://www.maa.org/past/putnam.html

Starting a Math Club

From Roy Thistle on September 4, 1996:

> For a couple years now, I have been trying to start a math club at our school... no luck. Perhaps a virtual club is the way to go? Some people might be thinking Usenet, but it is dying away because of the web. So there it is... lonely math guy... Does anyone have thoughts on how to stimulate some interest at my school or on how to start a virtual club?
>
> By the way, S.O.L.E. is an alternative school in Toronto with an excellent and dedicated staff who for the most part, unfortunately, are not interested in math. We have a great book club though... sigh... somehow it isn't enough.
>
> Thanks.

From Philip Spencer, University of Toronto on September 20, 1996:

An excellent question, and certainly if there's anything we can do here on the Mathematics Network to provide interest or help you start a virtual club, we'd be happy to do it.

Does anybody have any good thoughts or suggestions on how to stimulate interest in this kind of a situation?

From Alvin Mok, Richmond Hill High School on September 29, 1997:

I know exactly how you feel. I am trying to start a math club in my school. I even have my Math Department head helping me! But no luck. There is one thing you could do: try starting a math HELP club, that could attract a lot of people.

From Blanche Reyna on October 4, 1997:

I once taught at a junior high, about 4 years ago. I was successful at starting a math club and a Spanish club (at two schools). First of all, "math club" usually translates into "nerd" club; unfortunate, but true. So what I did was open both clubs to all students regardless of their GPA. Next I enticed students with goodies (e.g., picnics, movies). Immediately I had about 20 students. Afterwards, everyone wanted to join.

Later after the students realize that a "math club" is not a "nerd" club, we start playing chess, checkers, and slowly integrated math activities (e.g., district contests). The Spanish Club also started out the same way, I left the second year, but my principal informed me that the club was up and running 4 years later.

During the same year that I formed the Spanish Club, I approached the students with ideas of helping elders in the community, taking trips to San Antonio—multicultural activities—Mexican Cooking. It has been a great success.

The key to starting a club: entice students with something (you know, like the commercials on TV, like businesses do to get you into the door); remember teaching is a business, and students are our clients. Do not bombard students with academics at the beginning; do it slowly. It sure did work for me, and for both the math and Spanish Club. I hope I have been of help.

From Joseph, Hyde Park High School on January 24, 1998: Hi. I am the president of my school's math club but I really don't know what to do first. I really want this to be fun for the other students but I don't know where to begin.

Mathematical Communication

From Jennifer P on November 5, 1996:

> I am looking for information or resources on the topic of communications in math. If anyone knows where I can find such information please let me know.
>
> I can be emailed at Rossj@cybersensations.com

Teaching Linear Equations With CDROM Technology

From Allan Ast on December 13, 1996:

> Hi. I'm an adult education instructor for SIAST. I'm presently developing multimedia materials in the science area. Part of the program involves showing students how to solve linear equations using a CD ROM format. Does anyone have any suggestions as to how this could be presented using the potential of the CD ROM format. I've thought of using a manipulative approach similar to Alge-Tiles.
>
> Your suggestions would be greatly appreciated. Thanks!

From Myla V. de los Santos on July 28, 1997:

> I would like to find a better way of introducing word problems involving linear equations.

From Philip Spencer, University of Toronto on July 29, 1997:

> To be honest, I can't think of a better way of introducing linear equation word problems than simply stating them, choosing them from some real life application so that students can see the relevance. There are countless applications one can choose from. For example:
>
> You are renovating a house and have a leftover supply of wood trim: 17 long and 9 short. You want to use them up by making decorative window and door frames. A

window frame will use up 3 short pieces for the top and sides, and 1 long piece which is cut in two for a double-width sill. A door frame will use up a short piece for the top and two long pieces for the sides. How many window and door frames should you make, to completely use up your leftover trim? (Solve the system of linear equations $3x + y = 17$, $x + 2y = 9$).

A chemical company wants to produce 100 litres of oxygen and 50 litres of pure water. It does so by processing two types of raw material. Each litre of material A produces 0.6 litres of oxygen and 0.2 litres of water. Each litre of material B produces 0.3 litres of oxygen and 0.4 litres of water. How much of each raw material will be required to produce the desired quantities of oxygen and water? (Solve the system $0.6x + 0.3y = 100$, $0.2x + 0.4y = 50$).

These are just two off the top of my head. You might want to pick your own real-life situation that's of particular interest or relevance to your students.

Teaching Addition Using Fractions and Decimals Together

Asked by Blanca Reyna on October 3, 1997:

When showing students how to add, subtract, multiply or divide decimals, I I find it very useful to ask students to convert the decimals into fractional forms so they will see a connection to fractions, to money, to basic geometry. For example: $0.35+0.15 = 35/100 + 15/100 = 50/100 = 0.50$ which equals 50 cents in the context of dollars, or 1/2 the area of a unit square in the context of geometry.

Have any of you wonderful educators ever used this logic? If yes, what have your results been?

In great need for feedback, Blanche

Success in Mathematics and Future Success

Asked by Steve Bechtold, teacher, Blackduck on November 27, 1997:

I heard a blurb on the raio about 6 weeks ago about a study relating high school math success to future successes, like a much greater income. I thought I would read some published reprot but I have been unable to locate anything. Did anyone else hear this and get more details? For motivational reasons I am interested in this study, I think my students would work harder. thank you.

From Mr. Alexandru Pintilie, Bayview Glen School on February 12, 1998:

I have read a 2000 years old "published report" on this topic. One of Euclid's pupils complained that learning theorems was pointless - they were of no practical value. Euclid comanded a slave to give the boy a coin so he could make a profit from studying geometry.

Apart from its aesthetic value and its influence on one's capability to think, there is no benefit whatsoever in studying mathematics.

From Philip Spencer, University of Toronto on February 13, 1998:

But I would argue that those benefits you mention—especially the influence on one's capability to think—are enormous and do have a profound impact on a very wide variety of complex tasks, even ones that are not "mathematical" at first glance. I would not be surprised at all to hear that this translated into greater employability in today's world where many jobs require analytical skills (something that would not have been true in Euclid's day when most paid employment would have been manual in nature).

I am unacquainted with the study the first poster referred to. I do know, however, that many mathematicians in these days of government cutbacks to academic funding are finding *extremely* lucrative positions in industry, particularly in the financial world (mathematical finance is a very hot topic these days).

Student Misconceptions about Complex Numbers

Asked by Irem Yildirim, teacher on December 23, 1997:

What misconceptions and errors do students have in complex numbers?

Some misconceptions and errors are illustrated in the "fallacious proof that 1=2 using complex numbers", on our web site at http://www.math.toronto.edu/mathnet/falseProofs/second1eq2.html

For others, you may want to seek input from other high-school teachers. We've placed your question in the discussion area instead of the question area, in case anybody wants to add to it.

Motivating Students

From KF Lai on February 3, 1998:

I would like to seek for some creative ideas for the preparation of induction sets in the teaching of both Elementary Mathematics and Additional Mathematics of GCE O-Level.

Sometime, It is quite difficult to motivate students to study such topics as quadratic functions, functional relationship etc. I mean how to relate some of these topics with the students' real live or experiences.

As a teacher, I may say for example, the usefulness of integration technique in computing volume of an object with irregular shape, and that calculus is very important in the study of electronic engineering; but students may tell you that they are not interested to become an engineer, they would prefer to be a doctor instead!!

As a teacher, I cannot simply make such statement as "Well, boys and girls, whether you like it or not, you need to pass the exam!!"

Please kindly advise.

Sometimes, I am really exhausted of ideas on how to motivate them....

Sincerely from kflai@tm.net.my

Alor Setar, Kedah
West Malaysia