π in the Sky

The Pacific Institute for the Mathematical Sciences

http://www.pims.math.ca/pi

September 2003

ASTech RECIPIENT 2002

π in the Sky is a semi-annual publication of

The Pacific Institute for the Mathematical Sciences

PIMS is supported by the Natural Sciences and Engineering Research Council of Canada, the British Columbia Information, Science and Technology Agency, the Alberta Ministry of Innovation and Science, Simon Fraser University, the University of Alberta, the University of British Columbia, the University of Calgary, the University of Victoria, the University of Washington, the University of Northern British Columbia, and the University of Lethbridge.

This journal is devoted to cultivating mathematical reasoning and problem-solving skills and preparing students to face the challenges of the high-technology era.

Addresses:

π in the Sky	*π in the Sky*
Pacific Institute for	Pacific Institute for
the Mathematical Sciences	the Mathematical Sciences
449 Central Academic Bldg	1933 West Mall
University of Alberta	University of British Columbia
Edmonton, Alberta	Vancouver, B.C.
T6G 2G1, Canada	V6T 1Z2, Canada
Tel: (780) 492–4308	Tel: (604) 822–3922
Fax: (780) 492–1361	Fax: (604) 822–0883

E-mail: pi@pims.math.ca
http://www.pims.math.ca/pi

Contributions Welcome

π in the Sky accepts materials on any subject related to mathematics or its applications, including articles, problems, cartoons, statements, jokes, etc. Copyright of material submitted to the publisher and accepted for publication remains with the author, with the understanding that the publisher may reproduce it without royalty in print, electronic, and other forms. Submissions are subject to editorial revision.

We also welcome **Letters to the Editor** from teachers, students, parents, and anybody interested in math education (be sure to include your full name and phone number).

Cover Page: This picture was created for *π in the Sky* by Czech artist Gabriela Novakova. The scene depicted was inspired by the article on mathematical biology written by Jeremy Tatum, "Maths and Moths," that appears on page 5. Prof. Zmodtwo is again featured on the cover page, this time doing research on moths and butterflies.

CONTENTS:

This column is an open forum. We welcome opinions on all mathematical issues: research, education, and communication. Please feel free to write.

Reckoning and Reasoning

or

The Joy of Rote

by

by Klaus Hoechsmann[†]

You might have heard of this story, but it bears being repeated. In 1992, Lou D'Amore, a science teacher in the Toronto area, sprung a Grade 3 arithmetic test from 1932 on his Grade 9 class, and found that only 25% of his students could do all of the following questions.

1. **Subtract these numbers:** $9,864 - 5,947$
2. **Multiply:** 92×34
3. **Add the following:** $\$126.30 + \$265.12 + \$196.40$
4. **An airplane travels 360 kilometers in three hours. How far does it go in one hour?**
5. **If a pie is cut into sixths, how many pieces would there be?**
6. **William bought six oranges at 5 cents each and had 15 cents left over. How much had he at first?**
7. **Jane had $2.75. Mary had 95 cents more than Jane. How much did Jane and Mary have together?**
8. **A boy bought a bicycle for $21.50. He sold it for $23.75. Did he gain or lose and by how much?**
9. **Mary's mother bought a hat for $2.85. What was her change from $5?**
10. **There are 36 children in one room and 33 in the other room in Tom's school. How much will it cost to buy a crayon at 7 cents each for each child?**

This modest quiz quickly rose to fame as "The D'Amore Test." Other teachers tried it on their classes, with similar results. There was some improvement in Grades 10 to 12, where 27% of students could get through it, but they tend to be keener anyway since their less ambitious class-mates usually give up on quantitative science after Grade 9. All in all, the chance of acing the D'Amore Test appears to be independent of anything learned in high school.

At first glance this seems as it should be, because the test certainly contains no "high school material". On second thought, however, a strange asymmetry appears: while all students expect to use the first two R's (Readin' and Ritin') throughout their schooling and beyond, they drop the third R (Rithmetic) as soon as they can—if indeed they acquired it at all. Has it always been like this? I doubt it: my grandmother went to school only twice a week (being needed in yard and kitchen) but was later able to handle all the arithmetic in her little grocery store without prior attendance of remedial classes. She did not even have a cash register.

To many administrators, think-tankers, etc., this is beside the point, because we now live in the brave new computer age. A highly placed person who has likely never repaired a car engine, and probably knows little about computers, suggested that 20 years ago, "an auto mechanic needed to be good at working with his hands," whereas now he needs Algebra 11 and 12 to run his array of robots. For a more insights of this kind, you might wish to visit `www.geocities.com/Eureka/Plaza/2631/articles.html`, where electricians, machinists, tool-and-die makers, and plumbers are also included "among those who need Grade XI or XII algebra." It doesn't say what for.

Mechanics laugh at this: remember the breaker-point gaps, ignition timing, engine compression, battery charge, alternator voltage, headlight angle, and a multitude of other numerical values we had to juggle in our minds and check with fairly simple tools—today's gadgets make our jobs more routine, they say. But ministerial bureaucrats tend to believe the hype, with a fervour proportional to their distance from "Mathematics 12," which has gobbled up Algebra 12 in most places I know.

Aye, there's the rub: the third R has morphed into the notorious M. "What's in a name?," you ask, "that which we called rithmetic by any other word would sound as meek." How many times must you be told that M is hard and boring, and hear the refrain "I have never been good at M"? It is the perfect cop-out, acceptable even in the most exclusive company—a kind of egalitarian salute by which "normal" members of the species *homo sapiens* recognize one another. How can a teacher of, say, social studies be expected to develop vivid lessons around unemployment, national debt, or global warming—as long as these topics are mired in M? He/she still must mention numbers, to be sure, but can now present them in good conscience as disconnected facts, knowing that his/her students' minds will be uplifted in another class, by that lofty but (to him/her) impenetrable M.

Ask any marketing expert: labels are not value-free, they attract, repel, or leave you indifferent. Above all, they raise expectations, which, in the case of M, are as manifold and varied as the subject itself. Is it conceptualization, exploration, visualization, constructivism, higher-order thinking, problem solving—or all of the above? The guessing and experimenting goes on and on, producing bumper crops of learned papers and theses, conferences, surveys, and committees, as well as confused students and teachers. "This is the first time in history that Jewish children cannot learn arithmetic" said an

[†] **Klaus Hoechsmann** is a professor emeritus at the University of British Columbia in Vancouver, B.C. You can find more information about the author and other interesting articles at: `http://www.math.ubc.ca/~hoek/Teaching/teaching.html`.

Israeli colleague, referring to the state of Western style education in his country, where the recent Russian immigrants maintain a parallel school system.

Not every country has followed the R to M conversion. In the Netherlands and (what was) Yugoslavia, children still learn *rekenen* and *račun*, respectively, together with reading and writing. The more weighty M is left for later. Germany clung to *Rechnen* till the 1960's, and then rashly followed the American lead, pushing *Mathematik* all the way down to *Kindergarten*—with the effect of finding itself cheek-to-jowl with the US (near the end of the list) in international comparisons.

I hear the sound of daggers being honed: what is this guy trying to sell (in this culture we are all vendors), is it "Back to Basics"? Does he hanker for "Drill and Kill," for "Top Down" at a time when all good men and women aspire to "Bottom Up"? Readers unaccustomed to Educators' discourse might be puzzled at such extreme positions getting serious attention. They would immediately see middle ground between tyranny and anarchy, boot camp and nature trail, etc. Why do we always argue Black versus White? I really cannot explain it. Maybe it is because we need strident voices and must hold single notes as long as we can, in order to be noticed in this mighty chorus. How did we get here?

Although the benefits of planned obsolescence are obvious, they are not often mentioned to justify the present trend toward innumeracy. It is the relentless advance of technology which must be seen as the main reason for the retreat of archaic skills. Speech-recognizing computers already exist, and once they are mass-produced, writing will not need to be taught anymore, at least not at public expense. Whatever we now do with our hands and various other body-parts outside the brain will clearly fall into the domain of sports. Only in this spirit does it make sense to climb a mountain top that can be more safely reached by helicopter.

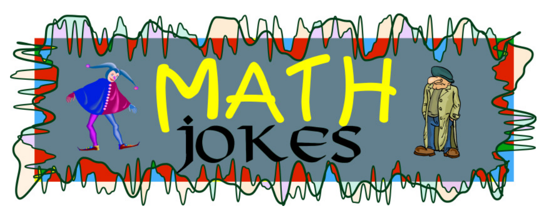
Before the advent of electric and later electronic calculators, computations had to follow rigid algorithms that allowed the boss or auditor to check them. This was "procedural knowledge" of an almost military kind—justly despised and rejected when it became obsolete. Oddly enough it did, however, have an important by-product: by sheer habit, *simple* calculations were done at lightning speed, and often mentally—of course with a large subconscious component. In many places, this "mental arithmetic" was even practised as a kind of sport, still "procedural," in some sense, but open to improvisation—more like soccer than like target shooting.

Look at the first question of the D'Amore Test: $9,864 - 5,947$. Abe did it the conventional way and had to "borrow" twice. Beth zeroed in on the last three digits, noting that 947 exceeded 864 by $36 + 47 = 83$, which she subtracted from 4000. Chris topped up the second number by 53 to 6000 and hence had to increase the first one to $9,864 + 53 = 9,917$. Dan and Edith had yet different ways, but all got $3,917$. On the second question, Abe again used the standard method, since he was a bit lazy but meticulous. Beth looked at the 92 and thought $100 - 10 + 2$, playing it very safe. Chris spotted one of his favourite short-cuts: $3 \times 17 = 51$, and reasoned that $9 \times 34 = 6 \times 51 = 306$, and so on. Dan was attracted to the fact that 92 was twice 46, which lies as far above 40 as 34 lies below it. Therefore 46×34 was $1600 - 36$, which had to be doubled to $3200 - 72$. Edith blurted out the answer 3128 and said she did not remember how she got it.

When I was in Grade 7, I knew such kids—and was irked by the fact that many played this mental game as well as they played soccer. Justice was restored when, in Grade 8, they were left in the dust by x and y but continued to outrun me on the playing field. Maybe they never missed the x and y in later life (unlike contemporary plumbers), but I am almost sure their "number sense" often came in handy. Today's kids are to acquire this virtue by doing brain-teasers and learning to "think like mathematicians," carefully avoiding "mindless rote."

Whenever I walk by the open door of a mathematician's work place, I see black or white boards covered with calculations and diagrams. How come they get to indulge in this "rote," while kids must fiddle with manipulations or puzzle till their heads ache? Could it be that we mathematicians sometimes engage in "mindful rote"—the kind known to musicians and athletes? If so, we ought to step out of the closet and tell the world about the *joy of rote*. Anyone who has observed young children will immediately know what we mean.

And while we're at it, we might reclaim ownership of the M-word, at least suggest that it be kept out of the K-4 world. This does not mean that schools should go back to teaching 'rithmetic—admittedly an awkward label. How about "reckoning and reasoning," a third and fourth R to balance the first two? They would be associated with good old common sense, and, as Descartes has pointed out, nobody ever complains of not having enough of that.



There are 10 kinds of mathematicians. Those who can think in binary and those who can't...

Two math professors are hanging out in a bar.

"You know," the first one complains. "Teaching mathematics nowadays is pearls for swine: the general public is completely clueless about what mathematics actually is."

"You're right!" says his colleague. "Look at the waitress. I'm sure she has no clue about any math she doesn't need to give out correct change—and maybe not even that."

"Well, let's have some fun and put her to the test," the first prof replies. He waves the waitress to their table and asks: "Excuse us, but you seem to be an intelligent young woman. Can you tell us what the square of $a + b$ is?"

The girl smiles: "That's easy: it's $a^2 + b^2$..."

The professors look at each another with a barely hidden smirk on their faces, when the waitress adds: "...provided that the field under consideration has characteristic two."

Q: What is the difference between a Ph.D. in mathematics and a large pizza?

A: A large pizza can feed a family of four...

A French mathematician's pick up line: "Voulez–vous Cauchy avec moi?"

Maths and Moths

Jeremy Tatum[†]

I don't reveal to many what some might regard as my somewhat eccentric hobby of rearing caterpillars and photographing the moths that ultimately emerge. This is my form of relaxation after the day is done, and my mind by then is usually far from mathematics.

Yet there is a moth, the Peppered Moth (*Biston betularia*), that lends itself well to mathematical analysis. It is common in Europe and in North America, including the west coast of Canada and the United States. It is often held to represent one of the fastest known examples of Darwinian evolution by variation and natural selection. A vast literature has accumulated on this moth, both by scientists and, I recently discovered, by creationists. The latter seek to disprove the hypothesis that it is an example of evolution, and their arguments do, I suppose, at least keep scientists on their toes to ensure that their evidence is compelling.

Figure 1: The normal "peppered" form of *Biston betularia*. Photographed by the author on Vancouver Island, British Columbia.

The normal form of the moth has a "peppered" appearance, shown by the specimen in Figure 1, which I photographed on Vancouver Island. When this normal form rests on a lichen-covered tree trunk it is very difficult to see; it is well protected by its cryptic coloration. There is another form that is almost completely black—the *melanic* form, illustrated in Figure 2 from a photograph taken in England by Ian Kimber. It is quite conspicuous when resting on a lichen-covered tree trunk, and it is at a grave selective disadvantage. The melanic forms are readily snapped up by hungry birds.

Figure 2: The melanic form of *Biston betularia*. Photographed by Ian Kimber in England.

In industrial areas of nineteenth century England, long before modern atmospheric pollution controls, factory chimneys belched out huge quantities of black smoke, which killed the lichens and coated the tree trunks with dirty black grime. Suddenly the "normal" form became conspicuous, and the melanic form cryptic. Within a few generations the populations of these moths changed from almost entirely "normal" to almost entirely "melanic." This is a situation that cries out for some sort of population growth analysis.

We first have to understand a little about genetics—and I hope that professional geneticists will forgive me if I simplify this just a little for the purpose of this article.

Figure 3: A melanic, a normal, and an intermediate form of *Biston betularia*. Photographed by Ian Kimber in England.

The colour of the moths' wings is determined by two genes, which I denote by M for melanic and n for normal. Each moth

[†] **Jeremy Tatum** is a former professor in the Department of Physics and Astronomy of the University of Victoria. His E-mail address is `universe@uvvm.uvic.ca`.

inherits one gene from each of its parents. Consequently the "genotype" of an individual moth can be one of three types: MM, Mn or nn. MM and nn are described as "homozygous," and Mn is "heterozygous." An MM moth is melanic in appearance, and an nn moth is normal. What does an Mn moth look like? Well, surprisingly, the heterozygous moth isn't intermediate in appearance; it is melanic. Because of this, we say that the M gene is dominant over the n gene; the n gene is recessive. That is why I have written M as a capital letter and n as a small letter.

(Actually the situation is rather more complicated than this, and there are indeed intermediate forms, as shown in Ian Kimber's photograph in Figure 3—but the purpose of this article is to illustrate some principles of mathematical analysis of natural selection, not to bog ourselves down in detail. So I'll keep the model simple and, to begin with, I'll suppose that just the two genes are involved and that one is completely dominant over the other.)

One can see now how vulnerable the M gene is in an unpolluted environment. Not only MM moths but also Mn moths are conspicuous and are easily snapped up by birds; the M gene doesn't stand a chance. But now blacken the tree trunks. MM and Mn are black and protected; the homozygous nn form is conspicuous. The n gene, however, does not disappear, because it is protected in the Mn individuals, which are black. The populations become predominantly composed of black individuals, some of which are MM and some are Mn.

Now for our first mathematics question. Suppose we have a large population, N, of moths. Each moth will have two genes that control wing colour, so there will be $2N$ such genes distributed among the N moths. Let us suppose that a fraction x of these genes are M and a fraction $1-x$ of them are n. What fraction of the population of moths will be genotypically MM, what fraction will be Mn, and what fraction will be nn? The answers are the successive terms of the expansion of $[x+(1-x)]^2$. That is, the fractions of MM, Mn, and nn moths in the population will be x^2, $2x(1-x)$, and $(1-x)^2$. (Verify that the sum of these is 1.) Since both MM and Mn are phenotypically melanic (i.e. melanic in external appearance), the fraction of melanic moths in the population will be $x(2-x)$ and the fraction of normal moths will be $(1-x)^2$.

Now, according to our theory, melanic moths in a polluted environment have a selective advantage over normal moths. Can we define "selective advantage" *quantitatively*? Let us suppose that a generation of moths emerges from their pupae such that the gene ratio $M:n$ is $x:1-x$, and hence that the genotype ratio $MM:Mn:nn$ is $x^2:2x(1-x):(1-x)^2$. Let us suppose that, by the time these moths are ready to lay their eggs to produce the next generation, the number of phenotypically melanic moths has been reduced by a factor α $(0\le\alpha\le1)$ and the number of phenotypically normal moths has been reduced by a factor γ $(0\le\gamma\le1)$. I define the selective *advantage* s of the melanic moths as

$$s=\frac{\alpha-\gamma}{\alpha+\gamma}. \tag{1}$$

This number lies between -1 to $+1$. If $s=-1$, the melanic form is at a severe disadvantage and indeed it is lethal to be black (as in unpolluted woods). No melanic moth will survive. If $s=+1$, the melanic form has a huge advantage; indeed it is lethal to be normal (as in polluted woods). No normal moths will survive. If $s=0$, neither form has an advantage over the other.

Note that both the MM and Mn moth numbers are reduced by the factor α. The next generation of moths, then, starts out with relative genotype frequencies

$$MM:Mn:nn=\alpha x^2:2\alpha x(1-x):\gamma(1-x)^2. \tag{2}$$

or, to normalize these proportions so that their sum is 1,

$$MM:Mn:nn=\frac{\alpha x^2}{\Sigma}:\frac{2\alpha x(1-x)}{\Sigma}:\frac{\gamma(1-x)^2}{\Sigma}, \tag{3}$$

where

$$\begin{aligned}\Sigma&=\alpha x^2+2\alpha x(1-x)+\gamma(1-x)^2\\&=(\gamma-\alpha)x^2-2(\gamma-\alpha)x+\gamma.\end{aligned} \tag{4}$$

Each MM moth contributes two M genes to the gene pool, and each Mn moth contributes one M gene. Therefore, the fraction of M genes in the new generation is $\frac{\alpha x^2}{\Sigma}+\frac{\alpha x(1-x)}{\Sigma}$, or $\frac{\alpha x}{\Sigma}$.

Now by inverting equation (1) we find that

$$\frac{\gamma}{\alpha}=\frac{1-s}{1+s}. \tag{5}$$

By using this, we can now express the gene frequencies in the new generation in terms of the selective advantage. Recall that in the initial generation the relative gene frequency was

$$M:n=x:1-x. \tag{6}$$

In the new generation it is

$$M:n=\frac{(1+s)x}{1-s+4sx-2sx^2}:\frac{1-s+(3s-1)x-2sx^2}{1-s+4sx-2sx^2}. \tag{7}$$

We can apply this to generation after generation to see how the proportion of M gene changes from generation in terms of the selective advantage (or disadvantage).

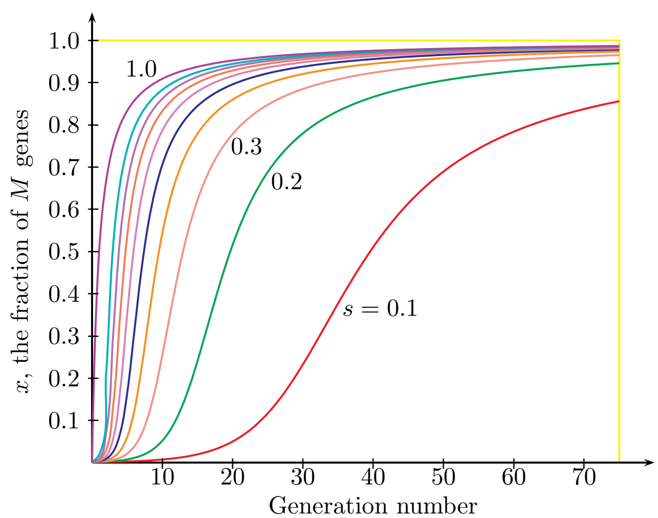


Figure 4: Complete dominance of M over n. Growth of the M-gene fraction x with generation number for ten selective advantages, from $s=0.1$ to 1.0 in steps of 0.1.

In Figure 4, I start with a fraction $x=0.001$ of M genes, and I watch the growth of this fraction with generation number for ten positive values of selective advantage—i.e. advantage to melanic moths on soot-covered tree trunks. Even for a mild advantage ($s = 0.1$), the fraction of M genes soon grows, while for a large advantage ($s = 0.9$) the growth of the M gene fraction is very rapid indeed, and this is believed to have happened to the moth *Biston betularia* in industrial areas in Britain. Note, however, that even if $s = 1.0$ (all normal phenotypes discovered and eaten by birds), the n gene survives (albeit in small numbers) because it is hidden and protected in the heterozygous Mn moths, which are phenotypically melanic.

What happens if we start with a high proportion of melanic genes, say $x = 0.999$, and put the moths in an unpolluted wood, where the tree trunks are lichen-covered, and the melanics are at a selective disadvantage (s is negative)? This in fact appears to be happening now in England, where air pollution controls are resulting in lichens recolonizing tree trunks that had become blackened with soot in a less environmentally-conscious era. Well, if we do the calculations, starting with $x = 0.999$, we find that almost nothing happens unless $s = -1$ exactly, in which case being a melanic phenotype is a death sentence, whether genotypically MM or Mn. The melanic gene is immediately extirpated. However, for any other negative value of selective advantage, very little happens for many generations, and the population remains predominantly melanic. This is because, even though the normal moths have the advantage, there are hardly any of them to enjoy it. Thus if the fraction of genes that are M is 0.999, the fraction of moths that are normal is only $(0.001)^2$, or 0.000001. For example, if we start with the fraction of M genes $x = 0.999$, and put them under a severe selective disadvantage of $s = -0.9$, even after 50 generations x is still 0.9927. However, after x has dropped to about 0.95, and normal (advantaged) moths begin to appear in the population in appreciable numbers, the decline of the M gene is rapid or even catastrophic. (Is this why the dinosaurs suddenly vanished after a long period of world dominance? Just a thought!) Indeed, since the disadvantaged M gene is not hidden and protected in the heterozygous moth, the M gene is eventually completely extirpated. In Figure 5, I have started with $x = 0.9$ (which is low enough for the start of rapid decline after a long period of quasistability), and we follow the decline of the M gene for a further 75 generations for 10 negative values of selective advantage.

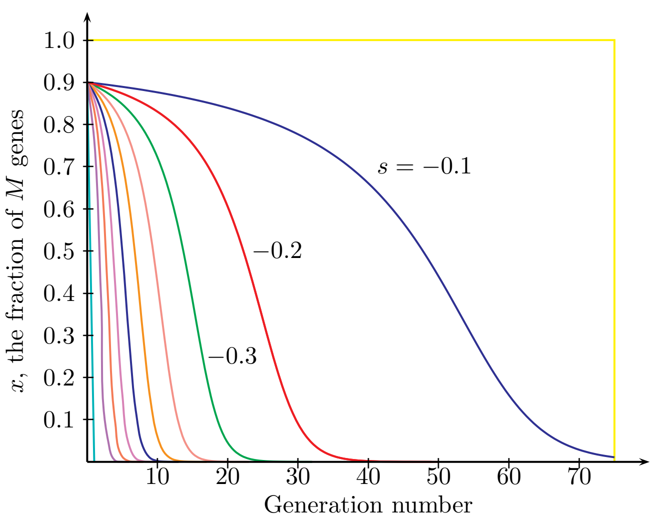


Figure 5: Complete dominance of M over n. Decline of the M-gene fraction x with generation number for ten selective advantages, $s = -0.1$, to -1.0 in steps of 0.1

So far, we have considered the case where one gene is completely dominant over another—but this is not always the case. In some species of moth the heterozygous form is intermediate in appearance to the two homozygous forms. In that case I'll use a small m for the "melanic" gene and a small n for the "normal" gene, so as not to give the impression that one is dominant over the other. The *three* possible genotypes are then mm, mn and nn, and they correspond to three phenotypes, melanic, intermediate and normal. I need to define selective advantage for each of the three forms, which I do as follows. I suppose that when one generation hatches from eggs, the relative numbers of genotypes in the population are in the proportion

$$mm : mn : nn = X : Y : Z. \tag{8}$$

Let us suppose that, by the time these moths lay their eggs to start the next generation, the numbers of melanic, intermediate and normal moths have been reduced by fractions α, β, and γ respectively. Then I define the selective advantages of the three forms as follows:

$$mm: \qquad s_1 = \frac{2\alpha - \beta - \gamma}{2\alpha + \beta + \gamma}, \tag{9}$$

$$mn: \qquad s_2 = \frac{2\beta - \gamma - \alpha}{2\beta + \gamma + \alpha}, \tag{10}$$

$$nn: \qquad s_3 = \frac{2\gamma - \alpha - \beta}{2\gamma + \alpha + \beta}, \tag{11}$$

These are not independent, and it takes a little algebra to show that they are related by

$$s_1 s_2 s_3 - 2(s_2 s_3 + s_3 s_1 + s_1 s_2) + 3(s_1 + s_2 + s_3) = 0. \tag{12}$$

They all have the property that they are in the range -1 to $+1$. A value of $+1$ means that the other two genotypes are completely destroyed, whereas a value of -1 means that that genotype is completely destroyed.

We can then do just what we did before when we went from equation (1) to equation (3). We suppose that the gene ratio of one generation is $x : 1 - x$. Then it works out that the fraction of m genes in the next generation is

$$\frac{(\alpha - \beta)x^2 + \beta x}{(\alpha - 2\beta + \gamma)x^2 + 2(\beta - \gamma)x + \gamma}. \tag{13}$$

Readers might like to convince themselves why it is not possible to invert equations (9)–(11) to express α, β, and γ uniquely in terms of the selective advantages, which is why it is more convenient and informative to write equation (13) in terms of α, β, and γ. One can then easily get a computer to apply this formula through generation after generation and see how the fraction of m genes changes with generation number.

There are four qualitatively different cases to consider.

I. The homozygous melanic mm is the fittest, and the homozygous normal nn is least fit. That is $\alpha > \beta > \gamma$. In Figure 6(a) I illustrate this for $\alpha = 0.4$, $\beta = 0.3$, and $\gamma = 0.1$. These correspond to $s_1 = +0.33\overline{3}$, $s_2 = +0.0\overline{90}$, and $s_3 =$ -0.55$\overline{5}$. I start with $x = 0.001$. The proportion of the m gene rapidly increases and the n gene eventually becomes extinct.

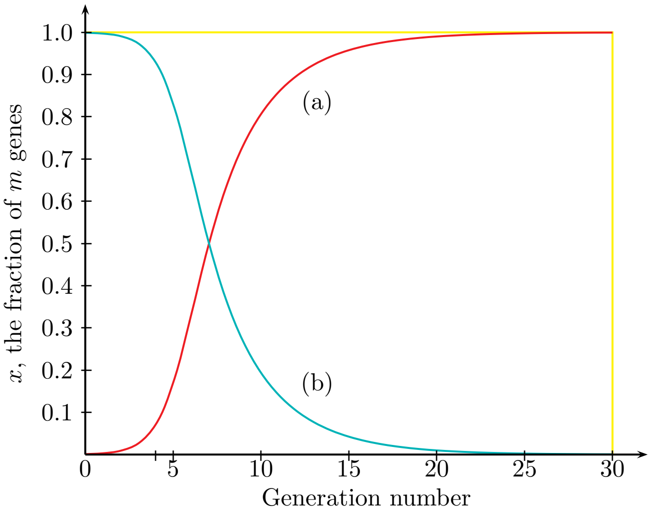


Figure 6: No dominance. (a) Case I. mm is the fittest, nn is the least fit. (b) Case II. mm is the least fit, nn is the fittest.

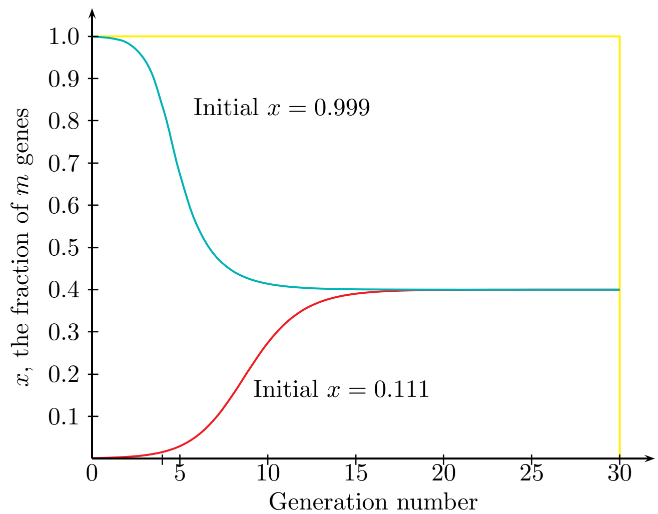


Figure 7: The heterozygous form has the advantage. Regardless of the initial m-gene fraction, high or low, an equilibrium situation ultimately results.

II. The homozygous melanic mm is least fit, and the homozygous normal nn is fittest. That is $\alpha < \beta < \gamma$. In Figure 6(b) I illustrate this for $\alpha = 0.1$, $\beta = 0.3$, and $\gamma = 0.4$. These correspond to $s_1 = -0.55\overline{5}$, $s_2 = +0.0\overline{90}$, and $s_3 = +0.33\overline{3}$. I start with $x = 0.999$. The proportion of the m gene rapidly decreases and eventually becomes extinct.

III. The heterozygous form has the advantage. That is $\beta > \alpha$ and $\beta > \gamma$. This case is rather more interesting! Regardless of the initial value of m-gene fraction x, the m-gene fraction eventually settles down to an equilibrium value x_e given by

$$x_e = \frac{\beta - \lambda}{2\beta - \alpha - \gamma}. \tag{14}$$

If the m-gene fraction is initially higher than this, it drops to the equilibrium value; if it is initially lower than this, it rises to the equilibrium value. This presumably means that if you have a population in which the three forms exist together for a long time, the heterozygous form is fitter than the other two. This case is illustrated in Figure 7, which I calculated for $\alpha = 0.2$, $\beta = 0.8$, and $\gamma = 0.4$. These correspond to $s_1 = -0.500$, $s_2 = +0.4\overline{54}$, and $s_3 = -0.11\overline{1}$. I started with $x = 0.001$ and $x = 0.999$.

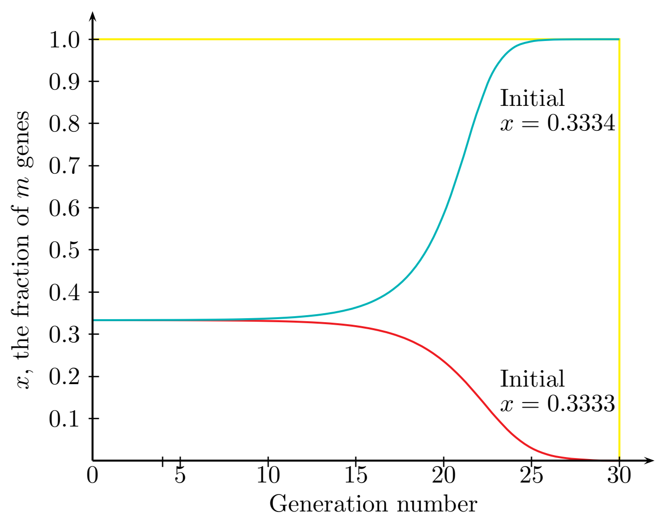


Figure 8: No dominance. The heterozygous form is at a disadvantage. The m-gene fraction goes to zero or infinity depending on whether its initial value is below or above a critical value.

IV. The heterozygous form is at a disadvantage. That is $\beta < \alpha$ and $\beta < \gamma$. Can you guess what will happen, just by thinking about it without actually doing the calculations? (Hint: Reverse the arrow of time!) What happens is that there is still an equilibrium m-gene fraction, and it is still given by equation (14)—but it is an *unstable* equilibrium! If the m-gene fraction starts ever so slightly above this equilibrium value, the fraction grows until the n-gene becomes extinct; and if the m-gene fraction starts ever so slightly below the equilibrium value, the m-gene becomes extinct. This case is illustrated in Figure 8, which I calculated for $\alpha = 0.8$, $\beta = 0.2$, and $\gamma = 0.5$. These correspond to $s_1 = +0.391$, $s_2 = +0.529$, and $s_3 = 0.000$. I started with $x = 0.3333$ and $x = 0.3334$. Very slight differences in initial conditions result in quite different outcomes.

Of course we have so far looked at some highly idealized situations. For example, we have assumed that the selective pressures remain constant generation after generation. If the environment changes at some time, this poses no particular difficulty: we can change the values of the selective advantage at any generation and resume the calculation with these new values. There is one example of biological significance that is also particularly amenable to this sort of mathematical calculation, and that is the study of mimicry. Some butterflies taste nasty and they are brightly coloured (warning coloration) so that birds can easily recognize them and leave them alone. Most tasty insects are cryptically coloured—difficult to find. But there are a few cheats. Some tasty insects mimic the bright colours of their horrible-tasting cousins; birds see the bright colours of the mimics and assume that they taste awful, so they leave them alone. This mimicry gives the cheat quite a selective advantage. But the cheat is effective only if the mimic is much rarer than the model. If the cheats are abundant, birds will not be taken in so easily and will soon unmask the fraud.

We can construct a plausible mathematical model of this situation. Let us suppose that there is a gene M for mimicry and a gene n for non-mimicry. To keep things simple, we'll suppose that the gene M is dominant over n (as you already guessed from the capital and small letters), so that there are just two forms of the insect—a mimetic form, which can be either MM or Mn, and a non-mimetic form. At some time, the fraction of mimetic insects in the population is X and the fraction of non-mimetic insects is $1 - X$. The selective advantage, we suppose, depends on the value of X, as we argued in the previous paragraph. Suppose, for example, that

$$s = -\frac{1}{2} + (1 - X)^{\frac{1}{2}}. \tag{15}$$

I have chosen this function quite arbitrarily, but it is at least plausible. It means that when the mimetic form is very rare (X very small), it has a distinct advantage ($s = +\frac{1}{2}$), and when it is common (X close to 1) it is at a decided disadvantage ($s = -\frac{1}{2}$). It has neither advantage nor disadvantage ($s = 0$) when $X = 0.75$. I admit that I also chose the function because it gives a very simple relation between s and x, the fraction of *genes* that are M. It is easy to show that

$$s = \frac{1}{2} - x. \tag{16}$$

Thus in terms of gene fraction (rather than mimetic insect fraction), s decreases linearly with x, going from $+\frac{1}{2}$ to $-\frac{1}{2}$, becoming zero for $x = \frac{1}{2}$. We can anticipate that, whatever the initial gene fraction, it will either increase or decrease until it reaches an equilibrium value of $+\frac{1}{2}$, when there is no selection but merely equal predation on mimetic and non-mimetic forms. The calculation is very easy. We just use equation (7) as before, but, instead of a constant value of s, we substitute $\frac{1}{2} - x$. The behaviour is illustrated in Figure 9, for initial gene fractions of 0.950 and 0.001.

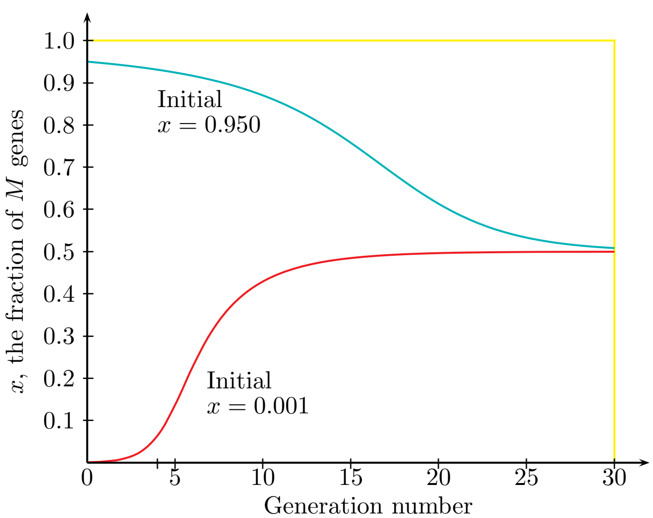
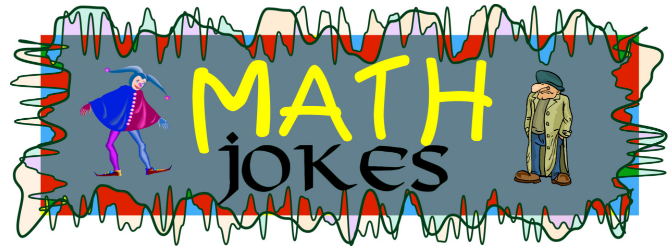


Figure 9: Complete dominance. The selective advantage of the melanic form depends upon its relative abundance in the population.

If you wish, you can further elaborate on these models of evolution by variation and natural selection and watch the changes in the population of the moths before your very eyes. I suppose it goes to show that, whatever subject happens to interest you, you will probably always find some application of mathematics to it that will make it even more interesting.



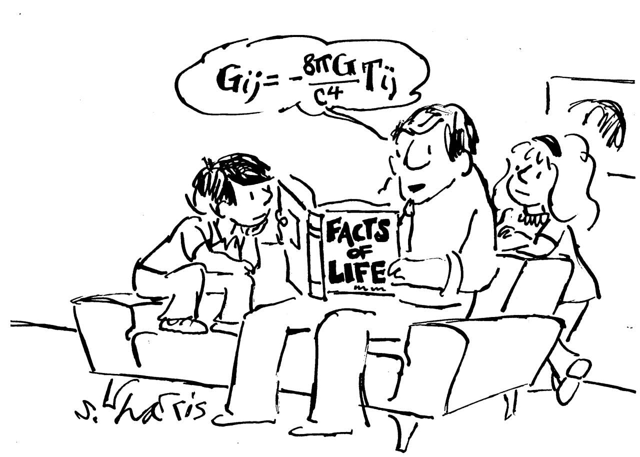
"My life is all arithmetic," the young businesswoman explains. "I try to add to my income, subtract from my weight, divide my time, and avoid multiplying..."

$$\lim_{8\to 9} \sqrt{8} = 3.$$

There are three kinds of mathematicians: those who can count to three, and those who can't...

Q: How can you tell that Harvard was planned by a mathematician?
A: The div school is right next to the grad school...

Mathematicians never die—they only lose some of their functions.



A mathematician named Haines
told—after wracking his brains—
that he had found
a new kind of sound
that travels much faster than planes.

Shouting Factorials!

Byron Schmuland†

h **Author's note:** This article may use ideas you haven't learned yet, and might seem overly complicated. It is not. But understanding Stirling's formula is not for the faint of heart; it requires concentrating on a sustained mathematical argument over several steps.

Even if you are not interested in all the details, I hope you will still glance through the article and find something to pique your curiosity. If you are interested in the details, but don't understand something, you are urged to pester your mathematics teacher for help.

Factorials!

Unbelievably large numbers are sometimes the answers to innocent looking questions. For instance, imagine that you are playing with an ordinary deck of 52 cards. As you shuffle and re-shuffle the deck you wonder: How many ways could the deck be shuffled? You reason that there are 52 choices for the first card, then 51 choices for the second card, then 50 for the third card, etc. This gives a total of

$$52 \times 51 \times 50 \times \cdots \times 2 \times 1$$

ways to shuffle a deck of cards. We call this number "52 factorial" and write it as the numeral 52 with an exclamation point: 52! This number turns out to be the 68-digit monster

80658175170943878571660636856403766975289505440883277824000000000000,

which means that if everyone on earth shuffled cards from now until the end of the universe, even at a rate of 1000 shuffles per second, we couldn't begin to see all the possible shuffles. Whew! No wonder we use exclamation marks!

For any positive integer n we calculate "n factorial" by multiplying together all integers up to and including n, that is, $n! = 1 \times 2 \times 3 \times \cdots \times n$. Here are some more examples of factorial numbers:

$1! = 1,\quad 2! = 2,\quad 3! = 6,\quad 4! = 24,$
$5! = 120,\quad 6! = 720,\quad 7! = 5040,\quad 8! = 40320,$
$9! = 362880,\quad 10! = 3628800.$

Stirling's Formula Factorials start off reasonably small, but by 10! we are already in the millions, and it doesn't take long until factorials are unwieldy behemoths like 52! above. Unfortunately there is no shortcut formula for $n!$, you have to do all of the multiplication. On the other hand, there is a famous approximate formula, named after the Scottish mathematician James Stirling (1692–1770), that gives a pretty accurate idea about the size of $n!$:

$$\text{Stirling's Formula: } n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

Before we continue, let's take a moment to contemplate the fact that n factorial involves nothing more sophisticated than ordinary multiplication of whole numbers, while Stirling's formula unexpectedly uses square roots, π (the area of a unit circle), and e (the base of the natural logarithm). Such are the surprises in store for students of mathematics.

Here is Stirling's approximation for the first ten factorial numbers:

$1! \approx 0.92,\quad 2! \approx 1.92,\quad 3! \approx 5.84,\quad 4! \approx 23.51,$
$5! \approx 118.02,\quad 6! \approx 710.08,\quad 7! \approx 4980.39,\quad 8! \approx 39902.39,$
$9! \approx 359536.87,\quad 10! \approx 3598695.62.$

You can see that the larger n gets, the better the approximation proportionally. In fact the approximation $1! \approx 0.92$ is accurate to 0.08, while $10! \approx 3598695.62$ is only accurate to about 30,000. But the proportional error for 1! is $(1!-.92)/1! = .0800$ while for 10! it is $(10!-3598695.62)/10! = .0083$, ten times smaller. This is the correct way to understand Stirling's formula, as n gets large, the proportional error $(n! - \sqrt{2\pi n}(n/e)^n)/n!$ goes to zero.

Developing approximate formulas is something of an art. You need to know when to be sloppy and when to be precise. We will make two attempts to understand Stirling's formula, the first uses easier ideas but only gives a sloppy version of the formula. We will follow that with a more sophisticated attack that uses knowledge of calculus and the natural log function. This will give us Stirling's formula up to a constant.

Attempt 1. To warm up, let's look at an approximation for the exponential function. The functions $1 + y$ and e^y have the same value and the same slope when $y = 0$, so that $1 + y \approx e^y$ when y is near zero (either positive or negative). Applying this approximation to x/n, for any x but with n much larger than x, gives $1 + x/n \approx e^{x/n}$. Now if we take the $(n-1)$st power on both sides, we get the approximation

$$\left(1+\frac{x}{n}\right)^{n-1} \approx e^{(n-1)x/n} \approx e^x.$$

Returning to factorials, we begin with an obvious upper bound. The number $n!$ is the product of n integers, none bigger than n, so that $n! \le n^n$. With a bit more care, we can write $n!$ precisely as a fraction of n^n as follows:

$$n! = \left(1-\frac{1}{2}\right)^1\left(1-\frac{1}{3}\right)^2\cdots\left(1-\frac{1}{n}\right)^{n-1} n^n.$$

I won't deprive you of the pleasure of working out the algebra to confirm that this formula is really correct. Using

† **Byron Schmuland** is a professor in the Department of Mathematical and Statistical Sciences at the University of Alberta. His E-mail address is schmu@stat.ualberta.ca. You can also visit his web page at http://www.stat.ualberta.ca/people/schmu/.

the approximation for the exponential function e^x we can replace each of the factors $(1-1/k)^{k-1}$ by e^{-1} and arrive at $n! \approx e\,(n/e)^n$. Because of cumulative errors, the formula $e\,(n/e)^n$ sorely underestimates $n!$, but it does have the right order of magnitude and explains where the factor "e" comes from.

Attempt 2. Our next attempt to get Stirling's formula uses the fact that, mathematically, addition is easier to handle than multiplication. Taking the natural log on both sides of $n! = 1 \times 2 \times \cdots \times n$, turns the multiplication into addition: $\ln(n!) = \ln(1) + \ln(2) + \cdots + \ln(n)$. As you might expect, our warmup problem this time is an approximate formula for the natural log function. We start with the series expansion

$$\frac{1}{2}\ln\left(\frac{1+x}{1-x}\right) = x + \frac{x^3}{3} + \frac{x^5}{5} + \frac{x^7}{7} + \cdots.$$

Substitute $x = 1/(2j+1)$ and rearrange to get

$$\begin{aligned}&\left(j+\frac{1}{2}\right)\ln\left(1+\frac{1}{j}\right) - 1\\ &= \frac{1}{3(2j+1)^2} + \frac{1}{5(2j+1)^4} + \frac{1}{7(2j+1)^6}\cdots.\end{aligned}$$

Now replacing the sequence of odd numbers $3, 5, 7, \ldots$ by the value 3 in the denominator makes the result bigger, so we have the inequality

$$\begin{aligned}&\left(j+\frac{1}{2}\right)\ln\left(1+\frac{1}{j}\right) - 1\\ &\le \frac{1}{3}\left(\frac{1}{(2j+1)^2} + \frac{1}{(2j+1)^4} + \frac{1}{(2j+1)^6} + \cdots\right).\end{aligned}$$

The sum on the right takes the form of the famous *geometric series*

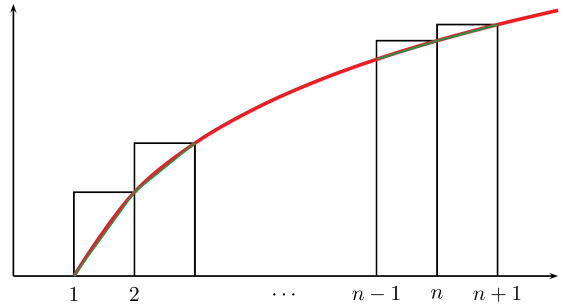
$$\rho + \rho^2 + \rho^3 + \cdots = \frac{\rho}{1-\rho}.$$

On making the replacement $\rho = 1/(2j+1)^2$, a little algebra yields

$$\begin{aligned}\left(j+\frac{1}{2}\right)\ln\left(1+\frac{1}{j}\right) - 1 &\le \frac{1}{3}\left[\frac{1}{(2j+1)^2-1}\right]\\ &= \frac{1}{12}\left(\frac{1}{j} - \frac{1}{j+1}\right).\end{aligned} \tag{1}$$

All that work was to show that $(j+1/2)\ln(1+1/j) - 1$ is pretty close to zero. If you are inclined, you could program your computer to calculate both sides of (1) for various values of j, just to check that the right hand side really is bigger than the left. Note that we have an upper bound in (1), instead of an approximate formula. This means that the values on the two sides are not necessarily close together, only that the value on the right is bigger.

You will be relieved to hear that we are finally ready to return to Stirling's approximation for $n!$. We want to approximate $\ln(n!) = \ln(1) + \ln(2) + \cdots + \ln(n)$, which is the area of the first $n-1$ rectangles pictured below. The curve in the picture is $\ln(x)$, and it reminds us that $\ln(1) = 0$.



The area of each rectangle is the area under the curve, plus the area of the triangle at the top, minus the overlap. In other words, using the definitions below we have $r_j = c_j + t_j - \varepsilon_j$.

$$\begin{aligned}\text{rectangle} &:= r_j = \ln(j+1),\\ \text{curve} &:= c_j = \int_j^{j+1} \ln(x)\,dx,\\ \text{triangle} &:= t_j = \frac{1}{2}[\ln(j+1) - \ln(j)],\\ \text{overlap} &:= \varepsilon_j = \left(j+\frac{1}{2}\right)\ln\left(1+\frac{1}{j}\right) - 1.\end{aligned}$$

The overlap ε_j is a small sliver shaped region that is barely visible in the picture, except in the first rectangle. Using the inequality (1) we worked so hard to establish, we add up on both sides and see that the infinite series satisfies $\sum_{j=n}^{\infty} \varepsilon_j < 1/(12n)$, for any $n = 1, 2, 3, \ldots$.

To approximate $\ln(n!) = \sum_{j=1}^{n-1} r_j$, we begin by splitting r_j into parts

$$\ln(n!) = \sum_{j=1}^{n-1} c_j + \sum_{j=1}^{n-1} t_j - \sum_{j=1}^{n-1} \varepsilon_j.$$

Since $\sum_{j=1}^{n-1} c_j$ is an integral over the range 1 to n, and $\sum_{j=1}^{n-1} t_j$ is a telescoping sum, this simplifies to

$$\begin{aligned}\ln(n!) &= \int_1^n \ln(x)\,dx + \frac{1}{2}\ln(n) - \sum_{j=1}^{n-1} \varepsilon_j\\ &= n\ln(n) - n + 1 + \frac{1}{2}\ln(n) - \left(\sum_{j=1}^{\infty} \varepsilon_j - \sum_{j=n}^{\infty} \varepsilon_j\right).\end{aligned}$$

Taking the exponential gives

$$n! = e^{1-\sum_{j=1}^{\infty}\varepsilon_j}\sqrt{n}\left(\frac{n}{e}\right)^n e^{\sum_{j=n}^{\infty}\varepsilon_j}.$$

Pause to note that this is an exact equation, not approximate. It gives $n!$ as the product of an unknown constant, the factor $\sqrt{n}\,(n/e)^n$, and a factor $e^{\sum_{j=n}^{\infty}\varepsilon_j}$ that converges to 1 as $n \to \infty$. The inequality $\sum_{j=n}^{\infty} \varepsilon_j < 1/(12n)$ then yields the bounds

$$C\sqrt{n}\left(\frac{n}{e}\right)^n \le n! \le C\sqrt{n}\left(\frac{n}{e}\right)^n e^{\frac{1}{12n}},$$

where $e^{\frac{11}{12}} \le C \le e$. Once we've identified $C = \sqrt{2\pi}$, we get

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n \le n! \le \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{1/12n}$$

If you've made it this far, congratulations! Now you see why Stirling's formula works. The part we skipped, to show that the unknown constant C is actually equal to $\sqrt{2\pi}$ is not that easy. But we've done enough hard work for today, so let's just accept this, and look at some other cool properties of the number $n!$.

Number of Digits For any $x > 0$ the formula $d(x) = \lfloor \log_{10}(x) \rfloor + 1$ gives the number of digits of x to the left of the decimal point. The funny looking $\lfloor\ \rfloor$ tells us to throw away the fractional part of the number. For moderate sized factorials we can simply plug this formula into a computer to see how many digits $n!$ has. For example, $d(52!) = 68$ and $d(1000000!) = 5565709$. But suppose we wanted to find the number of digits in a really large factorial, say googol factorial? (Googol means ten raised to the power 100 or 10^{100}). Even a computer can't calculate googol factorial, so we must use Stirling's formula. Let $g = 10^{100}$, substitute into Stirling's formula, and take log (base 10) on both sides to obtain

$$\begin{aligned}\log_{10}\left(\sqrt{2\pi g}\left(\frac{g}{e}\right)^g\right) &\le \log_{10}(g!)\\ &\le \log_{10}\left(\sqrt{2\pi g}\left(\frac{g}{e}\right)^g e^{1/12g}\right).\end{aligned} \tag{2}$$

Let's concentrate on the left-side $\log_{10}(\sqrt{2\pi g}(g/e)^g)$. Using the logarithm property and the fact that $\log_{10}(g) = 100$, we simplify this to $\log_{10}(\sqrt{2\pi}) + 50 + g(100 - \log_{10}(e))$. The hard part of this calculation is to find $\log_{10}(e)$ to over 100 decimal places, but the computer is happy to do it for us. Once this is accomplished we find that

$$\begin{aligned}\log_{10}\left(\sqrt{2\pi g}(g/e)^g\right) = {}&99565705518096748172348871081083\\ &39491770560299419633343388554621\\ &68341353507911292252707750506615\\ &682567.21202883\ldots.\end{aligned}$$

When we knock off the fractional part and add 1, we get $d(\sqrt{2\pi g}(g/e)^g)$. We can now find the number of digits in googol factorial by comparing with the upper bound. The right hand side $\log_{10}(\sqrt{2\pi g}(g/e)^g e^{1/12g})$ of (2) exceeds the left hand side only by the minuscule amount $\log_{10}(e^{1/12g}) = \log_{10}(e)/12g$. When this is added to the fractional part $0.21202883\ldots$, the first hundred or so digits after the decimal point are not affected. In other words, the three logarithms in (2) are so close together that knocking off the fractional part gives the same result. Therefore $d(\log_{10}(\sqrt{2\pi g}(g/e)^g e^{1/12g})) = d(\sqrt{2\pi g}(g/e)^g)$, and since $d(g!)$ is in between, it also must be the same.

Raising 10 to the power of the fractional part $0.21202883\ldots$ gives us the first few digits of $g!$, so we conclude that googol factorial is $g! = 16294 \cdots 00000$, where the dots stand in for the rest of the exactly

$$\begin{aligned}d(g!) = {}&99565705518096748172348871081083394917705602\\ &99419633343388554621683413535079112922527077\\ &50506615682568\end{aligned}$$

digits. This explains why no one can or ever will calculate all the digits of googol factorial. Where would you put it? A library filled with books containing nothing but digits? A trillion trillion computer hard drives? None of these puny containers could hold it. This super-monster has more *digits* than the number of atoms in the universe.

Trailing Zeros Looking back, you may notice that 52! ends with a stream of zeros. For that matter, all the factorials starting with 5!, have zeros at the end. Let's try to figure out how many zeros there will be at the end of $n!$. This doesn't rely on approximate values of $n!$ anymore, more importantly we need to understand the divisors of $n!$.

Each zero at the end of $n!$ comes from a factor of 10. For instance, 10! has two zeros at the end, one of which comes from multiplying the 2 and the 5.

$$\begin{aligned}10! &= 1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8 \times 9 \times 10\\ &= (1 \times 3 \times 4 \times 6 \times 7 \times 8 \times 9) \times (2 \times 5) \times 10\\ &= (36288) \times (100).\end{aligned}$$

The fact that 36288 is an even number means that there are extra factors of 2 that don't get matched with any 5's. Since there is always an excess of 2's, the number of trailing zeros in $n!$ is equal to the number of 5's that go into $n!$.

Imagine lining up all the numbers from 1 to n to be multiplied. You will notice that every fifth number contributes a factor of 5, so the total number of 5's that factor $n!$ should be about $n/5$. Since this isn't an integer, we knock off the fractional part and retain $\lfloor n/5 \rfloor$.

$$\begin{array}{cccccccccccccccccccccccc}1&2&3&4&5&6&7&8&9&10&11&12&13&14&15&16&17&18&19&20&21&\cdots&n\\ &&&&\uparrow&&&&&\uparrow&&&&&\uparrow&&&&&\uparrow&&&\end{array}$$

According to this formula, the number of trailing zeros in 10! is $\lfloor 10/5 \rfloor = 2$, and that checks out. But for 52! the formula gives $\lfloor 52/5 \rfloor = 10$, when there are really 12 trailing zeros. What's going on? The problem is that we forgot to take into account that the number 25 contributes *two* factors of 5, and does 50. That's where the extra two zeros come from.

Now we modify our formula for the number of trailing zeros in $n!$ to

$$z(n) = \lfloor n/5 \rfloor + \lfloor n/25 \rfloor + \lfloor n/125 \rfloor + \lfloor n/625 \rfloor + \cdots.$$

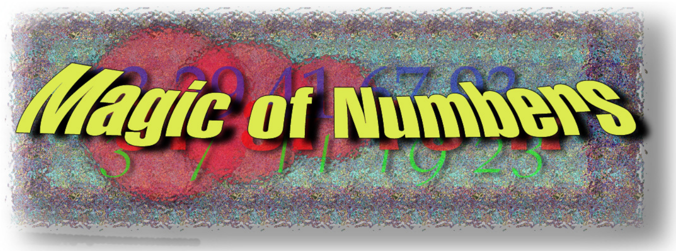
We have anticipated that all multiples of 125 give three factors of 5, multiples of 625 give four factors of 5, etc. Also note that if n is less than 25, for instance, then the formula $\lfloor n/25 \rfloor$ automatically returns a zero.

We can get an upper bound on the number of zeros by not knocking off the fractional part of $n/5^j$ and using the geometric series

$$z(n) = \sum_{j=1}^{\infty} \left\lfloor \frac{n}{5^j} \right\rfloor \le \sum_{j=1}^{\infty} \frac{n}{5^j} = \frac{n}{4}.$$

This turns out to be pretty close to the right answer. In other words, the number of trailing zeros in $n!$ is approximately $n/4$. For example, the number of trailing zeros in googol factorial works out to be exactly $z(g) = g/4 - 18$ or

$$\begin{aligned}&24999999999999999999999999999999999999999999999999\\ &99999999999999999999999999999999999999999999999982.\end{aligned}$$



A Generalization of Synthetic Division

Rohitha Goonatilake[†]

I. Introduction

In this article, we consider a procedure for division of polynomials. This is an alternative to a previously known process called long division of polynomials that involves the coefficients of polynomials. As we know, synthetic division works only for a divisor of the form $x - k$. In [1], Donnell showed that it can also be extended to a divisor of the form $x^n - k$ for $n \geq 2$. The purpose of this article is to extend this method to any polynomial divisor with a unit leading coefficient. This procedure has many applications; it is particularly important in factoring and finding the zeros of a polynomial. Several preliminary topics discussed in this article stem from [2].

Definition 1. Division Algorithm
If $p(x)$ and $d(x)$ are polynomials such that $d(x) \neq 0$, and the degree of $d(x)$ is less than or equal to the degree of $p(x)$, then there exists unique polynomials $q(x)$ and $r(x)$ such that

$$p(x) = d(x) \cdot q(x) + r(x),$$

where $r(x) = 0$ or the degree of $r(x)$ is less than the degree of $d(x)$. If the remainder $r(x)$ is zero, then we say that $d(x)$ divides evenly into $p(x)$. In this setting, $p(x)$, $d(x)$, $q(x)$, and $r(x)$ are respectively called *dividend, divisor, quotient,* and *remainder*.

Remark 1. The Division Algorithm can also be written as

$$\frac{p(x)}{d(x)} = q(x) + \frac{r(x)}{d(x)}.$$

The rational expression $p(x)/d(x)$ is called improper because the degree of $p(x)$ is greater than or equal to the degree of $d(x)$. On the other hand, the rational expression $r(x)/d(x)$ is called proper because the degree of $r(x)$ is less than the degree of $d(x)$. It is also assumed that $p(x)$ and $d(x)$ have no common factors.

II. Horner's Method

Horner's Method is a method of writing a polynomial in a nested manner. It gives us a method for evaluating polynomials that is very useful with a calculator. Consider the polynomial,

$$p(x) = 3x^3 + 8x^2 + 5x - 7.$$

Synthetic division by $(x - k)$ yields the following:

k	3	8	5	-7
	$-$	$3k$	$(3k+8)k$	$[(3k+8)k+5]k$
	3	$3k+8$	$(3k+8)k+5$	$[(3k+8)k+5]k-7$

Hence, by the remainder theorem, we know that $p(k) = [(3k+8)k+5]k - 7$. In terms of x, we can write

$$p(x) = 3x^3 + 8x^2 + 5x - 7 = [(3x+8)x+5]x - 7.$$

This is called Horner's method of writing a polynomial. It can be applied to any polynomial by successively factoring out x from each nonconstant term, as demonstrated in the following example.

$$\begin{aligned} p(x) &= 5x^4 - 3x^3 + x^2 - 8x + 7 \\ &= (5x^3 - 3x^2 + x - 8)x + 7 \quad \text{Factor } x \text{ from first four terms} \\ &= [(5x^2 - 3x + 1)x - 8]x + 7 \quad \text{Factor } x \text{ from first three terms} \\ &= \{[(5x - 3)x + 1]x - 8\}x + 7 \quad \text{Factor } x \text{ from first two terms} \end{aligned}$$

Before continuing the discussion of this topic any further, let us describe nested multiplication in a formal setting (so that it can be translated into a tableau), for a general polynomial $p(x)$ of degree m in Newton's form. It might be

$$p(x) = a_0 + a_1[(x - x_0)] + a_2[(x - x_0)(x - x_1)] + \cdots + a_m[(x - x_0)(x - x_1)\cdots(x - x_{m-1})].$$

This can be written succinctly as

$$p(x) = a_0 + \sum_{i=1}^{m} a_i \Big[\prod_{j=0}^{i-1} (x - x_j)\Big],$$

where the standard product notation has been used. The nested form of $p(x)$ is

$$\begin{aligned} p(x) =& a_0 + (x - x_0)\{a_1 + (x - x_1)[a_2 + \cdots + (x - x_{m-1})a_m]\cdots\} \\ =& \Big(\cdots\Big\{\big[a_m(x - x_{m-1}) + a_{m-1}\big](x - x_{m-2}) + a_{m-2}\Big\}\cdots \\ &+ a_1\Big)(x - x_0) + a_0. \end{aligned}$$

Thus, the polynomial $p(x)$ considered before, with all the x_js equal to zero, takes the nested form

$$p(x) = \Big\{\cdots\big[(a_m x + a_{m-1})x + a_{m-2}\big]\cdots + a_1\Big\}x + a_0,$$

with appropriate choices of $x_{m-1}, x_{m-2}, \cdots, x_0$ for a polynomial of degree m. For the tableau to be described in the next section, we write the divisor $d(x)$ of degree m with leading coefficient one in nested form as

$$d(x) = \Big(\cdots\Big\{\big[(x - k_1)x - k_2\big]x - k_3\Big\}x - \cdots - k_{m-1}\Big)x - k_m.$$

III. Synthetic Division

As illustrated above, there is a nice shortcut for long division of $p(x)$ by polynomials of the form $x - k$. The shortcut is called synthetic division and it involves the coefficients of the polynomial and k. The essential steps of this division tableau

[†] **Rohitha Goonatilake** is a professor in the Division of Mathematics, Department of Natural Sciences, Texas A&M International University, in Laredo, Texas 78041–1900, USA. His E-mail address is harag@tamiu.edu.

are performed by using only the coefficients. By the remainder theorem, we know that the remainder $r(x)|_{x=k} = p(k)$. Donnell [1] considered division by a polynomial of the form $x^n - k$ for $n \geq 2$. In the extension, we propose synthetic division of polynomials for any polynomial divisor. We insist that the leading coefficient of the divisor polynomial $d(x)$ be 1. In the event that the leading coefficient is different from 1, we divide both dividend $p(x)$ and divisor $d(x)$ by the leading coefficient of the divisor as required by this division tableau. It is understood that any divisions under consideration has this normalization and that the polynomials are written with descending powers of x. The latter is often referred to as *standard form.* We now describe the pattern for synthetic division of a given polynomial by any polynomial divisor using a carefully chosen set of worked examples. A proof of the general statement of this method is not attempted in this article due to the notational difficulties it may cause. In fact, such a proof is not within the scope of this article.

Suppose $p(x)$ is a polynomial of degree n, which is to be divided by a polynomial divisor $d(x)$ of degree m, where $1 \leq m \leq n$. This results in a quotient $q(x)$ of degree $n - m$ and a remainder $r(x)$ of degree $m - 1$ or less. The key steps of the procedure are explained below. Some discussions are very brief as we assume the reader is already familiar with the basic steps found in synthetic division and those of [2].

Step 1: The $n+1$ coefficients of $p(x)$ are arranged in order of descending powers of x in the top row of the division tableau. Zeros are used to replace any missing coefficients of the expansion.

Step 2: k_1 is chosen from the nested form. First it is placed outside to the left of the left extrema column. And for any k_j for $j \geq 2$, they are placed outside to the left of corresponding rows after the sum is computed for each and every column as j increases. Next, place j number of dashes (number of dashes correspond to the subscript of k) directly under first j number of coefficients of $p(x)$. For k_1, bring the first coefficient to be the first sum in the next row. The dashes are considered to have value 0 in computing this sum. Next, place $j+1$ dashes directly under the first $j+1$ coefficients of $p(x)$. Bring down the first $j+1$ coefficients to be the first $j+1$ sums in its row.

Step 3: The next step is to multiply each of these $j+1$ for $j \geq 1$ sum by k_{j+2}, placing the result of each multipliers in the next of $j+1$, for $j > 1$ positions to the right. Omit any product that would go beyond last column of the tableau.

Step 4: Add the next m (or possibly fewer) coefficients to these products and place sum in the next row. Continue this procedure until the bottom row is filled with $n+1$ numbers after all divisions are performed in this fashion for all k_j for $m \geq j \geq 1$. This can be easily understood by looking at the given set of examples.

Step 5: Finally, the first $n-m+1$ numbers on the last row from the left are chosen to be the coefficients of quotient polynomial $q(x)$. The next $m-1$ are related to coefficients of remainder polynomial $r(x)$. Before they are finally accepted as the coefficients of the remainder polynomial, they have to be adjusted. The numbers that appear in the boxes are the coefficients so obtained for the actual remainder polynomial.

Step 6: The number appearing next to the $\rightarrow$ is the product of the leading coefficients of the dividend and all the k_js for $j \geq 1$ in nested form of divisor $d(x)$. This is only done under limited situations as given in the next step. The sign $\downarrow$ means that the number is simply brought down to be the coefficient of the remainder. This is due to the fact that the sum is in question equals 0. The signs $\downarrow$ and $\rightarrow$ are used in accordance with the table given in Step 7 with priority given to $\downarrow$, whenever both occur.

Step 7: Depending on the number of k_js for $m \geq j \geq 1$, in the last row of the tableau, addends (of sums) in each row are subtracted from the number of the level of tableau. If the number of addends is 1, this subtraction is not carried out. If $j = 2$, this is done using the sum of the addend picked up from the top that amounts to $j - 1$. As j increases from 2 to 3, the sum of two addends is subtracted from the entries of the last row. This step is carried out for every column from the right. These numbers as well as their negated sums are underlined for easy referencing.

Example 1. Divide $2x^4 + 4x^3 - 5x^2 + 3x - 2$ by $x^2 + 2x - 3$. We write the divisor in a nested manner. Thus Horner's method applied to the divisor gives

$$x^2 + 2x - 3 = x(x+2) - 3 \quad \Rightarrow$$

Note $k_1 = -2$ and $k_2 = 3$.

$$\begin{array}{r|rrrrr}
-2 & 2 & 4 & -5 & 3 & -2 \\
 & - & -4 & 0 & 10 & \underline{-26} \\
\hline
3 & 2 & 0 & -5 & 13 & -28 \\
 & - & - & 6 & 0 & 3 \\
\hline
 & \underbrace{\underline{2}} & \underbrace{\underline{0}} & \underbrace{\underline{1}} & 13 & -25 \\
 & & & \rightarrow & -12 & \underline{26} \\
\hline
 & & & & \boxed{1} & \boxed{1}
\end{array}$$

The numbers above the braces give the coefficients of the quotient and those in boxes give the coefficients of the remainder polynomial. Now we have

Example	j for any products dropped	j - no. of products dropped at level j	$\rightarrow$ or $\downarrow$
1	2	$2-1=1$	$\rightarrow$
2	2	$2-0=2$	$\downarrow$
3	2	$2-1=1$	$\rightarrow$
4	2	$2-1=1$	$\rightarrow$
5	2, 3	$2-1=1$, $3-1=2$	use $\downarrow$ and not $\rightarrow$
6	2	$2-0=2$	$\downarrow$
7	2	$2-0=2$	$\downarrow$
8	2, 3	$2-1=1$, $3-1=2$	use $\downarrow$ and not $\rightarrow$
Remark 2	4	$4-2=2$	$\downarrow$

$$\frac{2x^4+4x^3-5x^2+3x-2}{x^2+2x-3} = 2x^2+0x+1+\frac{x+1}{x^2+2x-3}.$$

Example 2. Divide x^3-1 by x^2+x+1.
We write the divisor in a nested manner. Thus Horner's method applied to the divisor gives

$$x^2+x+1 = x(x+1)+1 \quad \Rightarrow$$

Note that $k_1=-1$ and $k_2=-1$.

$$\begin{array}{r|rrrr}
-1 & 1 & 0 & 0 & -1 \\
 & - & -1 & 1 & \underline{-1} \\
\hline
-1 & 1 & -1 & 1 & -2 \\
 & - & - & -1 & 1 \\
\hline
 & \underbrace{1} & \underbrace{-1} & 0 & -1 \\
 & & & \downarrow & \underline{1} \\
\hline
 & & & \boxed{0} & \boxed{0}
\end{array}$$

Now, similar to the previous example, we have

$$\frac{x^3-1}{x^2+x+1} = x-1+\frac{0x+0}{x^2+x+1}.$$

Example 3. Divide x^4+3x^2+1 by x^2-2x+3.
We write the divisor in a nested manner. Thus Horner's method applied to the divisor gives

$$x^2-2x+3 = x(x-2)+3 \quad \Rightarrow$$

Note that $k_1=2$ and $k_2=-3$.

$$\begin{array}{r|rrrrr}
2 & 1 & 0 & 3 & 0 & 1 \\
 & - & 2 & 4 & 14 & \underline{28} \\
\hline
-3 & 1 & 2 & 7 & 14 & 29 \\
 & - & - & -3 & -6 & -12 \\
\hline
 & \underbrace{1} & \underbrace{2} & \underbrace{4} & 8 & 17 \\
 & & & \rightarrow & -6 & \underline{-28} \\
\hline
 & & & & \boxed{2} & \boxed{-11}
\end{array}$$

Now, as before, we have

$$\frac{x^4+3x^2+1}{x^2-2x+3} = x^2+2x+4+\frac{2x-11}{x^2-2x+3}.$$

Example 4. Divide $x^4+x^3-x^2+2x$ by x^2+2x.
We write the divisor in a nested manner. Thus, Horner's method applied to the divisor gives

$$x^2+2x = x(x+2)+0 \quad \Rightarrow$$

Note that $k_1=-2$ and $k_2=0$.

$$\begin{array}{r|rrrrr}
-2 & 1 & 1 & -1 & 2 & 0 \\
 & - & -2 & 2 & -2 & \underline{0} \\
\hline
0 & 1 & -1 & 1 & 0 & 0 \\
 & - & - & 0 & 0 & 0 \\
\hline
 & \underbrace{1} & \underbrace{-1} & \underbrace{1} & 0 & 0 \\
 & & & \rightarrow & 0 & \underline{0} \\
\hline
 & & & & \boxed{0} & \boxed{0}
\end{array}$$

Now, as before, we have

$$\frac{x^4+x^3-x^2+2x}{x^2+2x} = x^2-x+1+\frac{0x+0}{x^2+2x}.$$

Example 5. Divide $x^4+3x^3-5x^2+6x+10$ by x^3+x^2+x+2.
We write the divisor in a nested manner. Thus Horner's method applied to the divisor gives

$$x^3+x^2+x+2 = ((x+1)x+1)x+2 \quad \Rightarrow$$

Note that $k_1=-1$, $k_2=-1$ and $k_3=-2$.

$$\begin{array}{r|rrrrr}
-1 & 1 & 3 & -5 & 6 & 10 \\
 & - & -1 & -2 & \underline{7} & \underline{-13} \\
\hline
-1 & 1 & 2 & -7 & 13 & -3 \\
 & - & - & -1 & -2 & \underline{8} \\
\hline
-2 & 1 & 2 & -8 & 11 & 5 \\
 & - & - & - & -2 & -4 \\
\hline
 & \underbrace{1} & \underbrace{2} & -8 & 9 & 1 \\
 & & & \downarrow & \underline{-7} & \underline{5} \\
\hline
 & & & \boxed{-8} & \boxed{2} & \boxed{6}
\end{array}$$

Now, as before, we have

$$\frac{x^4+3x^3-5x^2+6x+10}{x^3+x^2+x+2} = x+2+\frac{-8x^2+2x+6}{x^3+x^2+x+2}.$$

Example 6. Divide $2x^3-4x^2-15x+5$ by $(x-1)^2$.
We write the divisor in a nested manner. Thus Horner's method applied to the divisor gives

$$x^2-2x+1 = (x-2)x-1 \quad \Rightarrow$$

Note that $k_1=2$ and $k_2=-1$.

$$\begin{array}{r|rrrr}
2 & 2 & -4 & -15 & 5 \\
 & - & 4 & 0 & \underline{-30} \\
\hline
-1 & 2 & 0 & -15 & -25 \\
 & - & - & -2 & 0 \\
\hline
 & \underbrace{2} & \underbrace{0} & -17 & -25 \\
 & & & \downarrow & \underline{30} \\
\hline
 & & & \boxed{-17} & \boxed{5}
\end{array}$$

Now, as before, we have

$$\frac{2x^3-4x^2-15x+5}{(x-1)^2} = 2x+0+\frac{-17x+5}{x^2-2x+1}.$$

Example 7. Divide x^3 by x^2-x-1.
We write the divisor in a nested manner. Thus Horner's method applied to the divisor gives

$$x^2-x-1 = (x-1)x-1 \quad \Rightarrow$$

Note that $k_1=1$ and $k_2=1$.

$$\begin{array}{r|rrrr}
1 & 1 & 0 & 0 & 0 \\
 & - & 1 & 1 & \underline{1} \\
\hline
1 & 1 & 1 & 1 & 1 \\
 & - & - & 1 & 1 \\
\hline
 & \underbrace{1} & \underbrace{1} & 2 & 2 \\
 & & & \downarrow & \underline{-1} \\
\hline
 & & & \boxed{2} & \boxed{1}
\end{array}$$

Now, as before, we have

$$\frac{x^3}{x^2-x-1} = x+1+\frac{2x+1}{x^2-x-1}.$$

Example 8. Divide x^4 by $(x-1)^3$.

We write the divisor in a nested manner. Thus Horner's method applied to the divisor gives

$$(x-1)^3 = x^3 - 3x^2 + 3x - 1 = ((x-3)x+3)x - 1 \quad \Rightarrow$$

Note that $k_1 = 3$, $k_2 = -3$ and $k_3 = 1$.

$$
\begin{array}{r|ccccc}
3 & 1 & 0 & 0 & 0 & 0 \\
 & - & 3 & 9 & \underline{27} & \underline{81} \\
\hline
-3 & 1 & 3 & 9 & 27 & 81 \\
 & - & - & -3 & -9 & \underline{-18} \\
\hline
1 & 1 & 3 & 6 & 18 & 63 \\
 & - & - & - & 1 & 3 \\
\hline
 & \underbrace{1} & \underbrace{3} & 6 & 19 & 66 \\
 & & & \downarrow & \underline{-27} & \underline{-63} \\
\hline
 & & & \boxed{6} & \boxed{-8} & \boxed{3}
\end{array}
$$

Now, as before, we have

$$\frac{x^4}{(x-1)^3} = x + 3 + \frac{6x^2 - 8x + 3}{(x-1)^3}.$$

This tableau easily works to obtain coefficients of the quotient polynomial $q(x)$ and requires a set of additional steps prior to identifying the coefficients of remainder polynomial $r(x)$.

Remark 2. As we see below, the divisor of the form $x^n - k$, where $n \geq 2$, will require fewer steps and easily reduce to the techniques and related tableaus depicted in [2]. Hence, the results of this article generalize the standard synthetic division and its easy extension found in [2]. For example, let us apply our techniques to divide $x^5 + 4x^3 + 3x^2 - 2x + 8$ by $x^4 - 1$.

By Horner's method, we have $k_j = 0$ for $4 > j \geq 1$ and $k_4 = 1$. The division tableau is:

$$
\begin{array}{r|cccccc}
0 & 1 & 0 & 4 & 3 & -2 & 8 \\
 & - & 0 & 0 & \underline{0} & \underline{0} & \underline{0} \\
\hline
0 & 1 & 0 & 4 & 3 & -2 & 8 \\
 & - & - & 0 & 0 & \underline{0} & \underline{0} \\
\hline
0 & 1 & 0 & 4 & 3 & -2 & 8 \\
 & - & - & - & 0 & 0 & \underline{0} \\
\hline
1 & 1 & 0 & 4 & 3 & -2 & 8 \\
 & - & - & - & - & 1 & 0 \\
\hline
 & \underbrace{1} & \underbrace{0} & 4 & 3 & -1 & 8 \\
 & & & \downarrow & \underline{0} & \underline{0} & \underline{0} \\
\hline
 & & & \boxed{4} & \boxed{3} & \boxed{-1} & \boxed{8}
\end{array}
$$

As a result of all $k_j = 0$ for $m > j \geq 1$, the first three steps are not necessary and the number next to sign $\rightarrow$, (if any) is 0. But in this case the only missing entries in the last row are replaced by $\downarrow$ as indicated in the table. For $k_4 = 1$ and $k_j = 0$ for $3 \geq j \geq 1$, this reduces to

$$
\begin{array}{r|cccccc}
1 & 1 & 0 & 4 & 3 & -2 & 8 \\
 & - & - & - & - & 1 & 0 \\
\hline
 & \underbrace{1} & \underbrace{0} & \boxed{4} & \boxed{3} & \boxed{-1} & \boxed{8}
\end{array}
$$

Thus, by identifying coefficients of the quotient and the remainder polynomials, we have

$$\frac{x^5 + 4x^3 + 3x^2 - 2x + 8}{x^4 - 1} = x + 0 + \frac{4x^3 + 3x^2 - x + 8}{x^4 - 1}.$$

This topic has been presented in a course on college algebra. The method of synthetic division gives an easy way to perform the division of polynomials. The motivation is to use the method to obtain tableau for any polynomial division. This generalization of synthetic division can be easily learned, using the familiar steps of the usual method (tableau). All of the examples presented here, and many similar ones, can be done with a minimal amount of calculation. Thus, the method is a source of many nice exercises for undergraduate mathematics students for further exploration. As this saves time, an assignment or a project could be assigned to students to experiment with this procedure and to verify them by themselves using known results. Granted, this may not provide much motivation to students taking their first mathematics course, but the method is an interesting technique using key idea from the remainder theorem. As such, it is worthy of consideration for homework assignments, if not for formal inclusion in the course of college algebra.

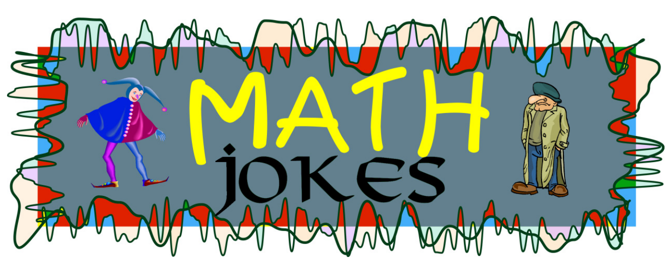
IV. Acknowledgment

The author wishes to thank Professor William Donnell [1] for presenting this topic at Texas A&M International University, Laredo, Texas for further exploration with some pertinent insights to the problem at hand.

References

1. William Donnell, *An Extension of Synthetic Division,* The Amatyc Review, Volume 18, Number 1, Fall 1996 published by the American Mathematical Association of Two-Year Colleges.

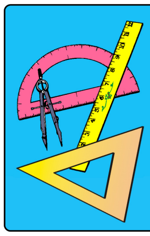
2. Roland E. Larson and Robert P. Hosteller, *College Algebra,* Second Edition, 1989, D. C. Heath and Company.



"EVEN THOUGH I'VE COME THIS FAR, I'M STILL NOT SURE WHICH IS THE NUMERATOR AND WHICH IS THE DENOMINATOR."

A mathematician is asked by a friend who is a devout Christian: "Do you believe in one God?"

He answers: "Yes—up to isomorphism."



Little Bits of Geometry

Why Not Use Ratios?

by Klaus Hoechsmann†

It was a year ago when last we met to spin this ongoing geometric yarn. On our way "From Rabbits to Roses," we had used a 17th century theorem by Girard Desargues—a dashing fellow judging by his picture—to convert a Golden Rectangle (Figure 1) into a Golden Triangle (Figure 2). The former is easy to construct, and the latter is the main ingredient of the pentagram, the mathematical backbone of roses, buttercups, cherry blossoms, etc.

Remember that a rectangle is *golden* if you are left with a smaller rectangle of the *same shape*—the upper grey strip in Figure 1—after removing a square (shown in lower part of that figure). From the big golden rectangle, we then made a triangle by rotating two of the longer sides inward until they met (as in Figure 2). To track down the angles so produced, we first turned the smaller golden rectangle through 90 degrees, so that it stood in the corner of the lower square, as shown. Having the same shape as its mother meant that its diagonal lined up with the maternal one.

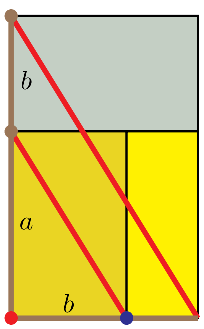


Figure 1

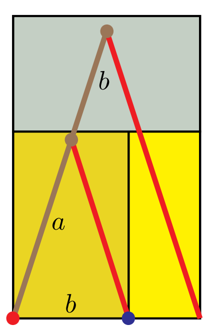


Figure 2

The main issue was to show that the slanted red lines remained parallel on the right since they started out that way on the left. That's were Desargues jumped in and helped, and for this installment of our story, the time seemed ripe to *prove* his theorem. This was the plan—until the folks from the Natural Ratio Association phoned in and scolded us for our lack of patriotism. "Why not use ratios?" they clamoured, hinting at boycott action unless we came up with a very good reason. "In both cases $(a+b) : a = a : b$, hence those triangles are similar—bang! No need for fancy French methods."

When asked what they meant by a "ratio," they would invariably suggest something like $5 : 9$ or $8 : 6$ or $37 : 31$. One of them admitted that it could get quite sticky like $1000001 : 7$, but the principle was always the same. "Like counting paving stones," he said. A few days later we received an envelope in the mail from a certain Moctezuma Ray in Waco, Texas. It contained a sheet of paper with a colourful drawing reminiscent of an Aztec head-dress as well as a long text, which began as follows.

Theorem: *Given a triangle ADE, let B and C be points on the line segments AD and AE, respectively. Then the lines BC and DE are parallel if and only if the ratios $AB : AD$ and $AC : AE$ are equal.*

The rest of the page contained a detailed proof, which we leave as an exercise for the reader. It was neatly divided into two parts: one for "if," the other for "only if." Both were impeccable—and tedious—in content, but attractively precise in lay-out.

"This man—oops! woman—evidently has not heard of incommensurability," said Karen, our history buff. "If this were written in Greek, I would date it around 600 B.C.—before the Pythagoreans, anyway." David, our trivia master, told Karen that Moctezuma was a man's name, and disagreed with the date on the grounds that they had no colour printers back then. "Let's frame it and call it *Ray's Theorem*," he suggested in a mocking tone, shaking his head in disbelief of Moctezuma's ignorance. Then he launched his browser to find the word "incommensurable" on the Internet, while I returned to my desk wondering how to explain it to our Texan correspondent.

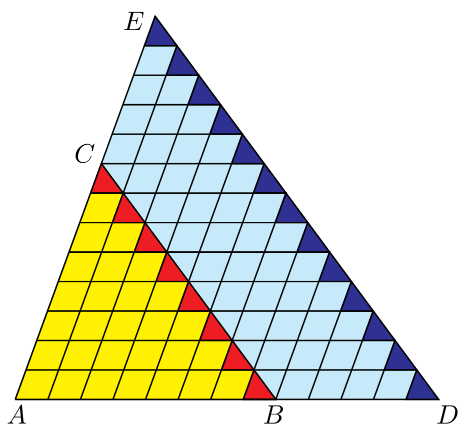


Figure 3

Basically, two lengths are "commensurable," if they can both be measured *exactly* with the same yardstick, in other words, if they can both be covered by a whole number of steps of the same size, like the line segments AB and AD in Figure 3. Actually, it is easier to describe it in two dimensions: a rectangle has commensurable sides if it can be completely paved with square tiles of the same size, like the rectangle below.

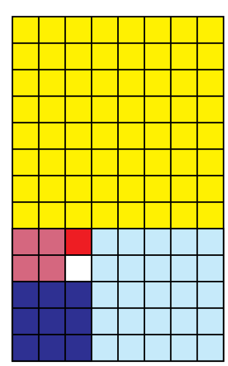


Figure 4

Practical people believe that this is always the case—if only you choose your units cleverly and small enough. Suppose your 2 by 4 foot steel plate turns out to be too wide by about 1/16 of an inch and too short by the same amount. Then you switch to millimeters and get 611 by 1218. If that is still not good enough for your boss, you can always go to micrometers, angstroms, or picofeet. The choices are endless: somewhere there is a unit that will somehow do the job.

This was the thinking even in scientific circles—until a

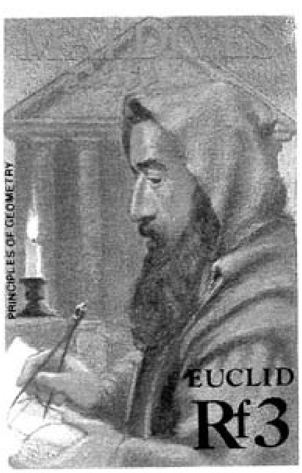
† **Klaus Hoechsmann** is a professor emeritus at the University of British Columbia in Vancouver, B.C. You can find more information about the author and other interesting articles at: http://www.math.ubc.ca/~hoek/Teaching/teaching.html

member of the Pythagorean sect found a pair of lengths that could never be commensurable. Most scholars say the first incommensurable pair discovered were the side and diagonal of a square, but some favour the sides of a golden rectangle.

Indeed, a golden rectangle could never be tiled as shown: otherwise, you could lop off the upper square and be left with a rectangle that is still tiled—with fewer tiles of the same size! This residual rectangle, however, is also golden—so you could repeat the process again and again. Even if you had started out with umpteen zillion tiles, these repeated reductions in their number would finally lead to a dead end. Thus, if you believe in perfect golden rectangles—remember, we constructed one last time—you have to believe in the existence of incommensurable magnitudes.

This realization struck the ancient scientists like a thunderbolt: all their arguments based on ratios were now invalid. Since no details about their immediate reaction are known, we have to rely on legends—and they run to extremes. At one end, we have the tale of a hundred oxen, which Pythagoras had slaughtered and roasted for a huge feast, to usher in the New Math. At the other end we have a nasty story of the hapless author of this "scandal" being pushed overboard to drown at sea. This is the version approved by the Natural Ratio Association—with the addition that the drowned wiseacre was a liar and a cheat.

Pythagoras died at about the time Socrates was born (around 400 BC), but the sect he founded outlived him by centuries. Based on kindness, truth, and contemplation—with a particular penchant toward mathematics—it was, of course, eventually persecuted and outlawed. In Plato's time, however, it was still flourishing, and one of its members, Archytas, was a respected philosopher, mathematician, and politician in Southern Italy. Today he is mainly remembered as the teacher of the brilliant Eudoxus—one of the most powerful minds in that era of mental giants. Though most of his energy was probably devoted to astronomy, Eudoxus touched on almost all branches of mathematics—and it was he who solved the puzzle of incommensurable ratios.



The Eudoxan theory of ratios takes up the entire Book V of Euclid's famous "Elements," and is applied to geometry from Book VI onward. The reader will forgive us for not reproducing these subtleties in the small space available here, but we cannot resist the temptation to lift the curtain a little bit. While the faces of Pythagoras and Eudoxus are hidden in the mists of the past, Euclid was close enough to the center of action—bustling Alexandria, Egypt, a scientific and cultural Mecca—that the image on the left may have some relation to reality. But more interesting than Euclid's face is his wily ingenuity. Instead of tackling Ray's Theorem head-on, he took a detour through a lemma that, at first glance, seems to have nothing to do with the theorem, but which is well-adapted to the ratios of Eudoxus.

In Book V, he had already proved that two magnitudes are in the same ratio to a third ($a : c = b : c$) if and only if they are equal ($a = b$). This is not as obvious as it sounds since the way Eudoxus uses the notion of ratio—which will be spelled out shortly—is more subtle than Mr. Ray's. But first let us see how Euclid argues his way to Ray's Theorem. From way back (Book I, 38 to 40) he knows: the lines BC and DE are parallel if and only if the blue-green area BDC equals the yellow-green area CBE. Now this appears in a new light: those two areas must be in the same ratio to the mauve area ABC. If he could change areas to lengths—i.e., claim

1. *the areas ABC and BDC are in the same ratio as the lengths AB and BD*

(of course, the same would hold for AC and CE), he could conclude that

2. *the lines BC and DE are parallel if and only AB is to BD as AC is to CE.*

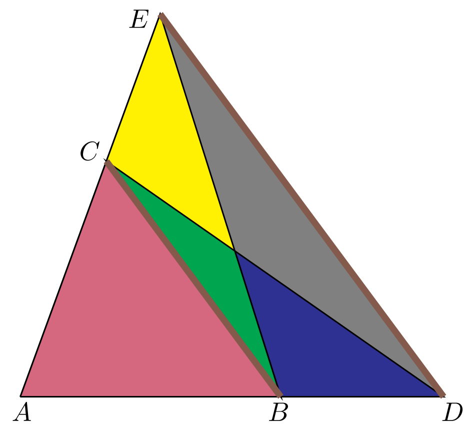


Figure 5

Roughly speaking, (1) says "triangular areas of the same height vary as their bases" and no sane person would doubt that. But Euclid's task is different: he has to derive it from first principles, where "vary" means "have the same ratios", and "same ratios" has a meaning prescribed by Eudoxus. Since names of points are arbitrary, let us ask why *the areas of CAB and ECB are in the same ratio as the lengths CA and EC*. Since D is gone, with all its kith and kin, we get a much cleaner picture than above—if, for a moment, you ignore all those extra triangles between P and Q. They are there because we're finally about to reveal how Eudoxus solved the riddle of incommensurable ratios. Note: for commensurable quantities a and b of whatever kind, you can find natural numbers n and m such that $na = mb$. For instance, if $a = CA$ and $b = EC$ in the picture from Waco, Moctezuma would count $n = 5$ and $m = 8$. However, for incommensurable quantities a and b, the equation $na = mb$ breaks down for any pair of natural numbers n and m. So Eudoxus says, in this case, that two ratios $a : b$ and $c : d$ should be considered equal (in modern notation, $a : b = c : d$) if, for all n

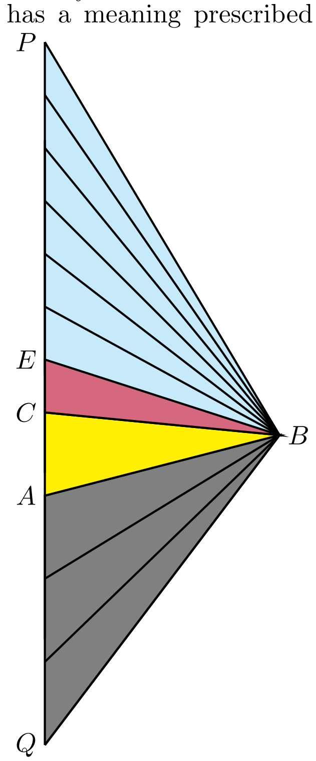


Figure 6

and m, these equations always break down the same way—in other words, "na exceeds mb" and "nc exceeds md" always go together. There are $m = 7$ triangles between C and P, all with the same base-length $b = EC$ and hence (by Book I, *ibidem*) all with the same area $d = ECB$; likewise, $n = 4$ triangles between C and Q with base-length $a = CA$ and equal areas $c = CAB$. For (1) to hold à la Eudoxus, mb should exceed na if and only if md exceeds nc—for *any* m and n, not just 7 and 4. In other words (check this carefully): PC should exceed CQ in length if and only if PCB exceeds CQB in area. But this is just the opening statement with a much coarser meaning: "triangular areas of the same height compare like their bases." This only asks whether one thing is bigger or smaller than another—not for a relation between the sizes. It is an easy extension of the same old results from Book I.

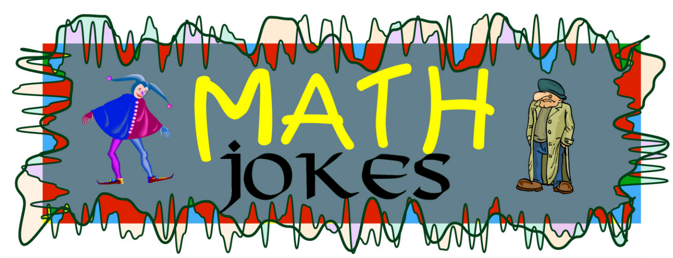
The attentive reader will have noticed that our (2) is not exactly the same as Ray's Theorem, which ties the parallelity of those lines to the equality $AB : AD = AC : AE$, not (as we have it) $AB : BD = AC : CE$. But Euclid can parry that objection: in Book V, he had already proved that $a : b = c : d$ is the same as $a : (a+b) = c : (c+d)$. Since $AD = AB + BD$ and $AE = AC + CE$, this makes our theorem just a minor variation of Ray's.

As I was wondering how to say this as smoothly as possible, David burst out of his office and exclaimed: "Now I get it, incommensurable ratios are just irrational numbers!" He was visibly delighted by this insight, and slapped his forehead with his palm: "I should have twigged on to this as soon as I saw that picture from Waco." He further suggested that the Natural Ratio Association should call itself the Rational Number Association, but I pointed out that the acronym RNA was already taken by ribonucleic acid. "I know you must have good reasons for speaking in riddles, but please listen to the way I understand it and tell me where I am wrong," he said, as he held up Moctezuma's picture. Then he explained that since A,B, and D lay on the same line, there would be a *vector equation* $AD = sAB$, for some real number s, and similarly $AE = tAC$, for some real number t. If $s = t$, we could subtract the second equation from the first and get $DE = sBC$, hence parallelity of DE and BC. On the other hand, if we *chose* an E' such that $AE' = sAB$, we'd have DE' parallel to BC, as before, and—since there is only one line parallel to BC through D—E' and E would be the same, and so would s and t.

I was impressed. "What's wrong with that?" he demanded to know. Then Karen—probably the source and inspiration of this argument—emerged from their shared office to reinforce the question. Nothing was wrong, I had to admit, except that we were hanging out in different mental worlds. "You're putting Descartes before Dehorce," I said. They did not appreciate that feeble pun. "Is Dehorce the French name for Euclid?" Karen asked and tried to laugh. But she did have the right hunch of what was meant. They had been working in the *Cartesian* plane—so called because it was invented by René Descartes, a contemporary of our hero Desargues, though much more famous and always on the run from displeased potentates (echoes of Pythagoras). The Cartesian plane consists of "points" that are nothing other than pairs (x, y) of numbers. You can make 3, 4, 5, or higher dimensional "spaces" from the same ingredients; so the only thing that's really real is your initial pool of numbers—"real" numbers, in the case at hand.

I offered to show how they related to the ratios of Eudoxus: "If na exceeds mb, then certainly $10na$ exceeds $10mb$, but maybe $7na$ will be enough, or even fewer. That's how you get the next term in your decimal expansion." "Not *that* again," David groaned, "infinite decimal expansions racing—wham!—right through the Milky Way and beyond, skewering every galaxy in their path—running on eternally, while we are supposed to think of them as quietly sitting on the number line. In school I just accepted this as a figure of speech, but the more we discuss it now, the flakier it seems." Nor was he won over by Cauchy sequences, nested intervals, Dedekind cuts, and the like: we had been through that before. "It's the infinity I don't understand. I can talk about such things—even real numbers, as I just did—but I am not really thinking what I am saying." He paused for a moment: "Well, I do have some kind of mental image, but not a precise one—except in the *commensurable* case.... Whoopee!—at least now I understand the issue. I should sign up with the NRA," upon which he disappeared into his office.

"Would you call that a lack of imagination?" I asked Karen. "Maybe rather an excess of honesty," she answered, "David is one of those people who won't call a spade a spade, unless he sees a spade."



©Copyright 2003
Sidney Harris

Q: Why do mathematicians never rob banks?

A: They have financial mathematics to do it for them.

Math Discovery

It's All for the Best:

How Looking for the Best Explanations Revealed the Properties of Light

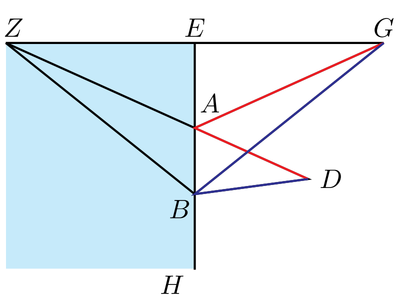
by

Judith V. Grabiner[†]

People who do mathematics and physics are always looking for the best: what is the shortest distance, the quickest path, the system using the least energy? Why do we do this? More precisely, how have we scientific types come to believe that mathematical laws describing the world often maximize something, or minimize something? I'm going to tell you the story of how this came to be. And it begins in the first century of the Common Era, with Heron of Alexandria. You may know him as the author of Heron's formula for the area of a triangle, or as the inventor of a little steam engine shaped like a rotating lawn sprinkler. But he did something much more important.

Heron seems to have been the first to use maximal and minimal principles to explain something in physics. What he explained was the equal-angle law of reflection of light. "Everybody thinks," he says, "that light travels in straight lines." (The Greeks, by the way, thought sight went from the eye to the object.) How fast is light? Well, we see the stars as soon as we look up, even though the distance is infinite, so the rays go infinitely fast. "Well," Heron says, "all fast-moving objects travel in straight lines." Why? To get where they're going faster. By reason of its speed, the object tends to move over the shortest path. And that's a straight line.

OK, that's ordinary unimpeded light. What is the law of reflected light? It was already known that light is reflected at equal angles. But Heron asks, "why?" "Same reason," he says! "Of all possible incident rays from a given point reflected to a given point, the shortest path is the one that is reflected at equal angles." Here's his diagram and proof: (Catoptrics, 4)



Heron concludes that it is in conformity with reason that light takes the shortest path. He has deduced the law of reflection rationally, from a principle of economy.

Consider AB a plane mirror, G the eye, and D the object of vision. Let a ray GA be incident upon this mirror such that $\angle EAG = \angle BAD$. Let another ray GB also be incident upon the mirror. Draw BD. I say that

$$GA + AD < GB + BD.$$

Draw GE from G perpendicular to AB, and extend GE and AD until they meet, say at Z. Draw ZB. Now, we know

$$\angle BAD = \angle EAG, \quad \text{and} \quad \angle ZAE = \angle BAD$$

(as vertical angles). Therefore, $\angle ZAE = \angle EAG$. And since the angles at E are right angles,

$$ZA = GA \quad \text{and} \quad ZB = GB.$$

But $ZD < ZB + BD$ and $ZD = ZA + AD$. We deduce that $GA + AD < GB + BD$.

Let us turn now to another ancient mathematician, Pappus of Alexandria, about 300 C.E., who was interested in problems like this: of all plane figures with the same perimeter, which has the greatest area? Pappus introduces his discussion of isoperimetric problems by enlisting an unusual mathematical ally: the honeybee. Pappus says, "God gave men the best and most perfect notion of wisdom in general and of mathematical science in particular, but a partial share in these things he allotted to some of the unreasoning animals as well." "Now," he says, "how do bees set up their honeycombs?" We observe that they keep their honey in clean and pure ways, and that they divide their combs into hexagons. Why hexagons? They have contrived this by virtue of a certain geometrical forethought—the figures must be contiguous to one another—their sides common, so that no foreign matter could enter the interstices between them and so defile the purity of their produce. "Only three regular polygons," Pappus says, "are capable by themselves of exactly filling up the space about the same point: the square, the equilateral triangle, and the hexagon." The hexagon has a larger area than the square or the triangle with the same perimeter, so the bees conclude that the hexagon will hold more honey for the same expenditure of wax used for the perimeters. Thus the bees have solved what we today call the problem of economically tiling the plane with regular polygons. "Since we are smarter than bees," Pappus continues, "we'll go on and investigate more general problems of this type." For instance, he says that the circle has the greatest area for figures of the same perimeter, and the sphere has the largest volume for solids with the same surface area. These results were proved in 1884 by H. A. Schwarz.

One more maximal principle comes from the Greeks, this one cosmological. Why are there so many different kinds of things in the universe? Here's Plato's answer: the universe contains the maximal amount of being. The historian Arthur Lovejoy calls this the principle of plenitude. "The amount of being in the universe is maximal," says Plato, "because of the goodness and lack of envy of the creator. The creator exists, so he makes the universe as much like himself as possible: full of existing things." This principle was picked up by various theologians and philosophers and was unbelievably influential. For instance, it is the seventeenth-century argument for the infinite universe and for the proposition that all stars have inhabited planets.

I could read to you hundreds of later statements influenced by Plato, Heron, and Pappus about maximum or minimum principles and economy. Here are four: (1) Olympiodorus, 6th century: "nature does nothing superfluous or any unnecessary work." (2) Robert Grosseteste, 13th century: "nature always acts in the mathematically shortest and best possible way." (3) William of Occam, 14th century, in the doctrine now known as Occam's razor: "the simplest explanations are the best." And, (4) in the Renaissance, Leonardo da Vinci: "nature is economical and her economy is quantitative. For instance, living things eat each other so that the maximum amount of life can exist from the minimum amount of material."

[†] **Judith V. Grabiner** is Flora Sanborn Pitzer Professor of Mathematics at Pitzer College in Claremont, California, U.S.A. Her E-mail address is jgrabiner@pitzer.edu. You can also visit her web site at http://www.pitzer.edu/academics/faculty/grabiner/.

All these traditions find their culmination in the 17th century explanation of the refraction of light by Pierre de Fermat. First, recall the *law of refraction*:

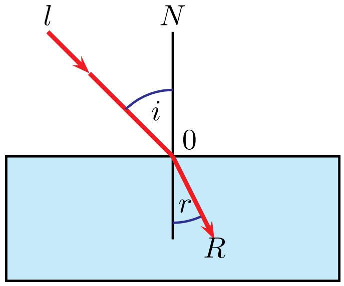


Figure 1: The angles of incidence and refraction.

$$\frac{\sin i}{\sin r} = \text{constant},$$

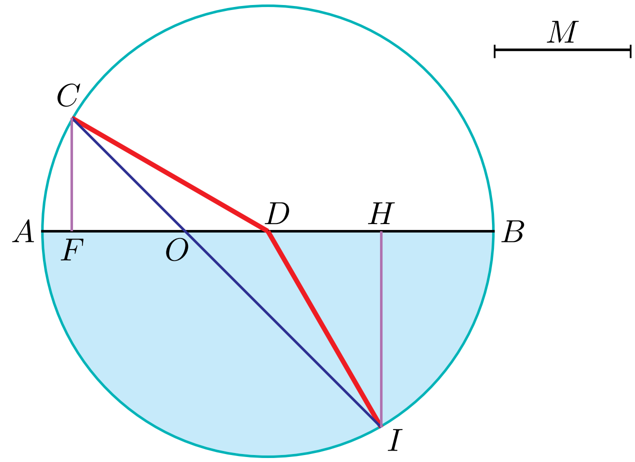
what we now call Snell's law, independently discovered by Descartes.

Let's look at Fermat for a minute. As you may know, Fermat anticipated much that later became calculus, including working out a general method of maxima and minima. As we read both in calculus textbooks and in physics textbooks, Fermat's principle in optics says that when light is refracted from one medium to another, it takes the path that minimizes the time.

But Fermat was a mathematician, not a physicist. Between inventing analytic geometry, methods of maxima and minima, tangents, areas, famous work in number theory, and—with Pascal—inventing probability theory, not to mention making a living as a lawyer, he didn't do any other physics. So how did he get involved in optics?

To begin with, Fermat and Descartes independently invented both analytic geometry and methods of finding tangents. Descartes, when he heard about Fermat's work, responded with disrespect. Descartes claimed that Fermat's tangent method wasn't general (it was in fact better than Descartes'), and that Fermat should read Descartes' Geometry to learn what was what. Fermat was annoyed. When he read Descartes' work on optics, he was in no mood to be charitable. He strongly criticized Descartes' derivation of Snell's law. Descartes imagined light being a mechanical motion of particles in an ether. "When a ray of light crosses a boundary from one area to another where the ether has different density," Descartes said, "it was like when a ball hits a tennis net. The component of velocity parallel to the net is unchanged, but the one perpendicular is changed since it can't penetrate the net." You'd think the light would be slowed down; that's what Fermat thought too. But no, according to Descartes' work, coming into the denser medium, it gets a little kick from the net. So it's bent toward the perpendicular.

Fermat thought Descartes' justification was nonsense, so he attacked the problem himself. Fermat's approach was motivated by a guy who is hardly a household name: Marin Cureau de la Chambre, who in 1657 wrote about the law of reflection exactly the way Heron had. He said that nature always acts along the shortest paths. But by path, Fermat did not mean distance. Instead, Fermat used an idea of Aristotle's, that velocity in a medium varies inversely as the medium's resistance to motion. So for Fermat, the path to be minimized in refraction is not the sum of the two lines CD



and DI, but a sum involving multiples of those lines, the multiples being determined by the ratio M of the resistances. He constructs I so that the sum $CD + DI \cdot M$ is minimal, where CD and DI are the lengths minimizing this quantity. Minimizing this path in fact minimizes the time traveled, although Fermat, wanting to avoid committing himself to any position about the yet-unmeasured speed of light, doesn't say so. In 1662, Fermat applied his own method of maxima and minima to find conditions on the fastest path. He expected to derive the true law of refraction—and was astounded that it was the same Descartes/Snell law! Well, that's how things are. But now we know why this is the law: light follows the fastest path. And Fermat's explanation of why this must be the law of refraction helped establish it as an important physical law.

So we see that explaining physical phenomena is explaining, "Why is it like this?" Showing that laws maximize or minimize something long predates calculus. But of course calculus makes it much easier and much more natural, as Leibniz realized. So let's look at his ideas on the subject.

Leibniz was both a philosopher, who wanted to maximize and minimize things, and an inventor of the calculus that allowed him do it very well. In fact, Leibniz called his differential calculus a new method of maxima and minima. And Leibniz's first non-trivial application of his new calculus was to derive the law of refraction.

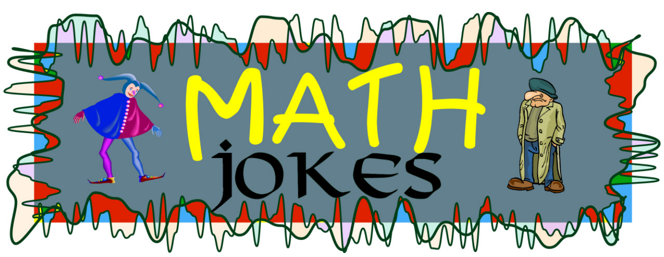
Why, for Leibniz, should light follow the shortest path? Not just because Fermat said so. It is an example of something much more general in Leibniz's philosophy. Leibniz's first principle is the principle of sufficient reason. Nothing happens without a reason. "For instance," he says, "Archimedes used the principle of sufficient reason to show that a balance with equal weights at equal distances must not incline to either side. It leads naturally to symmetry and economy." But Leibniz said more: "Every true proposition [about nature] has an *a priori* proof; a reason can be given for every truth. The first decree of God, [is] to do always what is most perfect." Leibniz also used Plato's Principle of Plenitude: "How many beings must this world contain? ALL possible kinds." For Leibniz, this is the best of all possible worlds. What does he mean by this? The best of all possible worlds is that in which the quantity of existence is as great as possible. (Voltaire made fun of him for saying this, but so what?) The divine will chooses the best world, the one with the greatest number of things in it, and it is precisely for that reason that the laws of nature are as simple as possible; that way, God can find room for the most possible things. Leibniz says, "if God had made use of other, less simple laws, it would be like constructing a building of round stones, which leave more space unoccupied than that which they fill." Leibniz's philosophical justification for the simplest laws, the maximum existence, the shortest paths, made people even more excited about finding such laws.

Back to our original question: HOW did we come to learn to search for and find these explanations? Philosophical ideas about God as a rational economist, powerfully reinforced by examples from geometrical optics and the geometrical insights of honeybees, and vastly accelerated by the techniques of calculus, have, in the centuries after Leibniz, led to the discovery of curves of quickest descent, the principle of least action in physics, to the calculus of variations, and to philosophical ideas—agree with them or not as you choose—like the idea of the greatest good for the greatest number and Adam Smith's free-market principle that individuals striving to maximize their profit leads to the most efficient organization of the entire economy.

I think that the successful search for the best and most economical solution helped reinforce the idea of progress in science. It also strengthens and reinforces the idea so embedded in our teaching and practice that we cannot imagine that it was ever otherwise: the idea that mathematics in general, and calculus in particular, is the best way to model this most mathematically elegant of all possible worlds.

Principal Sources

1. M. R. Cohen and I. E. Drabkin, *Source Book in Greek Science*.
2. Arthur Lovejoy, *The Great Chain of Being*.
3. Michael Mahoney, *The Mathematical Career of Pierre de Fermat*.
4. Dirk J. Struik, *Source Book in Mathematics, 1200–1800*.



As the waters receded and Noah's ark finally came to rest on top of Mount Ararat, Noah and his family, along with all the animals, left the ark.

But after forty days below deck on an overcrowded boat, none of the animals was in the mood for mating, and Noah worried about how to repopulate the Earth.

So, he tore down one of the ark's masts, cut it into pieces, and built a table out of the logs. Then he told one of the snakes to perform a lascivious dance on top of the table, while all of the other animals gathered around it. After a while the snake's seductive moves showed an effect: one animal after another started to sway in the rhythm of the snake's dance. They began to sneak away in pairs until the dancing snake and her mate were finally left all alone. They too disappeared, leaving Noah and his family overjoyed that the animal population would soon be back on track.

Q: What does this story from the book of Genesis teach us about math?

A: If you want to go forth and multiply, all you need are a log table and an adder!

A mathematician and his best friend, an engineer, attended a public lecture on geometry in thirteen-dimensional space.

"How did you like it?" the mathematician wanted to know after the talk.

"My head's spinning," the engineer confessed. "How can you develop any intuition for thirteen-dimensional space?"

"Well, it's not even difficult. All I do is visualize the situation in arbitrary N-dimensional space and then set $N = 13$."

"Students nowadays are so clueless," the math professor complains to a colleague. "Yesterday, a student came to my office and wanted to know if General Calculus was a Roman war hero..."

A math professor accepts a new position at a university in another city and has to move. He and his wife pack all their belongings into cardboard boxes and have them shipped off to their new home. To sort out some family matters, the wife stays behind for a few more days while her husband leaves for their new residence. The boxes arrive before the wife rejoins her husband. When they talk on the phone in the evening, she asks him to count the boxes, just to make sure the movers haven't lost any of them.

"Thirty nine boxes altogether," says the prof on the phone.

"That can't be," the wife exclaims. "The movers picked up forty boxes at our old place."

The prof counts once again, but again his count only reaches 39. The next morning, the wife calls the moving company and complains. The company promises to check; a few hours later, someone calls back and reports that all forty boxes did arrive. In the evening, when the prof and his wife are on the phone again, she asks: "I don't understand it. When you count, you get 39, and when they do, they get 40. That's more than strange..."

"Well", the prof says. "This is a cordless phone, so you can stay on the line and count with me: zero, one, two, three...."

Q: Do you know any catchy anagram of Banach–Tarksi?

A: Banach–Tarksi Banach–Tarski....

Q: How can you tell when you are dealing with the Mathematics Mafia?

A: They make you an offer that you can't understand.

"YES, MATH ANXIETY IN A MATHEMATICIAN IS A BIT UNUSUAL."

A. N. Kolmogorov and His Creative Life

Alexander Melnikov[†]

In the famous article "*The Architecture of Mathematics*" (1948) by Nikolas Bourbaki, there is a note that states with regret that there isn't any mathematician, even among those having the broadest erudition, who is not a stranger in some areas of the vast World of Mathematics. **Andrei Nikolaevich Kolmogorov (April 25, 1903–October 20, 1987)**[*], being the foremost mathematician of the 20th century, brings us a real counterexample to this assertion.

His scientific horizon covered almost every area of mathematics. His unique insight and deep understanding resulted in more than 300 research papers, monographs, and various textbooks. His creative activities, impregnated by fundamental ideas and outstanding results, initiated several completely new areas of mathematical investigation.

The list of the areas affected by his contributions includes the theory of sets, trigonometric and orthogonal series, measure and integration theory, mathematical logic, topology and homology theory, celestial mechanics, approximation theory, turbulence, ergodic theory, superposition of functions, information theory, functional analysis and above all—probability theory, which was transformed by Kolmogorov into real mathematical science. However, his interests were not limited only to mathematics. He also exhibited interests in areas of applications to biology, geophysics, statistical control of production, ballistics theory, and even the theory of poetry, where his originality and penetrating thoughts made permanent impact.

His collected works are divided into two volumes (*Mathematics and Mechanics*[1] and *Theory of Probability and Mathematical Statistics*[2]). These two titles reflect the fact that our vast mathematical world is divided into two parts, which can be classified as the deterministic and random phenomena kingdoms. Kolmogorov was like a trailblazer in both kingdoms. He discovered many unexplored regions and filled them with new exciting ideas. He put forward an ambitious program for a simultaneous and parallel study of the complexity of deterministic phenomena and the uncertainty of random phenomena, which practically dominated his whole life. The full value of his work is still being realized and explored today.

A.N. Kolmogorov was born on April 25, 1903 in the town Tambov in Russia. His father—Nikolai Kataev, a son of a clergyman working as an agronomist, and his mother—Mariya Yakovlevna Kolmogorova, were never married. He was named after his grandfather, Yakov Stepanovich Kolmogorov instead of his own father. The mother tragically died in childbirth at Kolmogorov's birth, which happened while she was travelling back home from Crimea. To make things worst, Kolmogorov's father practically abandoned his child and was never involved in his upbringing. The sister of his mother, Vera Yakovlevna Kolmogorova, took the responsibility for his care. This educated and free thinking woman, whom A.N. Kolmogorov always treated as his real mother, passed to her nephew an independence of opinion, the desire to understand rather than memorize, a disapproval of laziness, a despise of poorly performed tasks, a high sense of responsibility, and the aspiration to face difficult challenges. The fact that Kolmogorov's family originated from nobles, caused additional complications in the years following, due to the Russian revolution.

A.N. Kolmogorov spent his youth in the family estate in Tunoshna. After finishing school, he worked briefly as a conductor on the railway. During his teenage years, besides mathematics, his interests included Russian history. He enrolled at Moscow University in the autumn of 1920. At that time mathematics was not his greatest passion. He studied various other subjects including metallurgy and in particular history. He even wrote a serious treatise on the 15th century history of the Russian city Novogrod. One of the well-known anecdotes describes Kolmogorov's history teacher explaining to him that "maybe in mathematics one proof is considered to be sufficient, however in history it is preferable to have at least ten proofs!"

The first creative period in the Kolmogorov's life as a mathematician was greatly influenced by his teachers: Professor Stepanov, who directed a seminar on trigonometric series, and Professor Lusin—his supervisor. At that time, in 1922, when Kolmogorov was just a 19-year-old undergraduate student, he discovered the famous example of a Fourier series that is almost everywhere divergent. This truly surprising result still highlights the depth of his ideas and his profound geometrical intuition. Nowadays, every textbook on the theory of trigonometric and orthogonal series is bound to include Kolmogorov's example.

Kolmogorov's interests in probability theory originated in 1924, and since then his authority has been considered to be of the greatest importance in that branch of mathematics. In 1924–1928, he succeeded in finding necessary and sufficient conditions for the convergence of series of independent random variables and the law of large numbers, one of the main statements of classical probability theory. He graduated in 1925, but persisted to stay at Moscow University for four more years as a "research student." However, he was forced to conclude his studies when stricter rules regulating the duration of the enrollment at the university were introduced in 1929. Kolmogorov's difficulties finding a new place to research were resolved by Aleksandrov, who secured for him a vacant position in the Institute of Mathematics and

[†] **Alexander Melnikov** is a professor in the Department of Mathematical and Statistical Sciences at the University of Alberta. His web site is http://www.math.ualberta.ca/Melnikov_A.html and his E-mail address is melnikov@ualberta.ca.

[*] This article is written to commemorate the 100th anniversary of A.N. Kolmogorov.

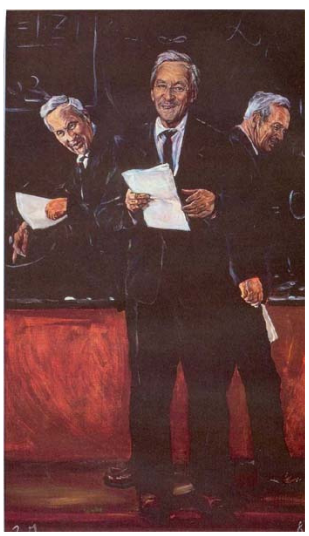
[1] Published by Nauka, Moscow, in 1985.

[2] Published by Nauka, Moscow, 1986. English translation of the Kolmogorov's collected works was published by Kluwer Publishing House.

Mechanics at Moscow University.

During the years 1929–1933, Kolmogorov worked on measure theory with the purpose of establishing a solid basis for probability theory, which resulted in Kolmogorov's classical monograph "*Fundamental Concepts of Probability Theory.*" This book settled not only new directions in the development of probability theory as a branch of mathematics (which was one of the famous *Hilbert's Problems* presented at the World Congress of the Mathematicians in 1900), but it also laid the foundations for the creation of the theory of random processes. Other famous concepts of Kolmogorov in this area were presented in his remarkable paper "*Analytical Methods in Probability Theory*" (1931). This work paved the road to the modern theory of Markov processes. The story tells that some of this research was done during a boat trip on the Volga river during the summer of 1929. Kolmogorov and Alexandrov rented a small boat and camping equipment from the "Society for Proletarian Tourism and Excursions," which was created for the purpose of promoting active living among workers in the Soviet Union. During this trip they covered about 1300 kilometers staying in secluded idyllic surroundings, sunbathing, swimming and doing mathematics.

In 1931 Kolmogorov was hired as a professor at Moscow University, and from 1937 he held the chair of theory of probability. Kolmogorov always maintained a very active lifestyle that included skiing, rowing, and long excursions on foot—on average about 30 kilometers. He loved swimming in the river, especially in the early spring when the snow and ice was just beginning to melt. His physical fitness was matched by his enormous productivity. During the decade preceding the Second World War, Kolmogorov published more than sixty papers on probability theory, mathematical statistics, topology, projective geometry, theory of functions, mathematical logic, and mathematical biology.



Painting of Kolmogorov by his former student Dmitrii Gordeev

It was during this period that Kolmogorov made significant contributions to homology theory. He also constructed an example of an open map of a compact set onto a compact set of higher dimension. His attention was also attracted by the mechanics of turbulence, that is, the irregular pulsations of velocity, pressure, and other hydrodynamical quantities occurring in flows of fluids or gases. Kolmogorov developed a rigorous statistical approach to provide a mathematical description of such flows and in 1941 formulated his most famous (and still unproven) conjecture in this area, known as the *Two-Thirds Law*. It states that *in every turbulent flow, the mean square difference of the velocities at two points a distance r apart is proportional to* $r^{2/3}$.

His interests touched every branch of science. He wrote about the growth of crystals, astronomy, and even genetics. One of his research papers brought him into a confrontation with academician T.D. Lysenko. Lysenko denied the existence of genes, claiming that evolution occurred because organisms inherit characteristics that have been adapted by their ancestors. Lysenko's theory was denounced by the scientific community as completely wrong. Kolmogorov, armed with scientific evidence courageously opposed Lysenko's views and supported Mendel's theory.

The post-war period in Kolmogorov's scientific life can be characterized by two words: harmony in diversity. Kolmogorov was working on an unusually large spectrum of topics: probability theory, classical mechanics, ergodic theory, the theory of functions, information theory, and algorithm theory. For him, these subjects, seemingly remote and unrelated, were all interconnected by completely unexpected links. This characteristic of Kolmogorov's understanding is perfectly illustrated by his remarkable works, written in the 1950's, on the theory of dynamical systems. He was motivated by the problem of three and more bodies, going back to Newton and Laplace. In particular, this problem is related to explaining the observations of the so-called quasiperiodic motions of small planets.

Kolmogorov in his Moscow office

Kolmogorov solved this important problem for most of the initial conditions. In following decades, further application of his theory made it possible to solve a variety of other problems. Later, the method of Kolmogorov was improved by Arnold and Moser and now is known as the KAM-theory.

In 1955, Kolmogorov's interests turned to information theory and subsequently, to the 13th Hilbert problem, which postulated that certain continuous functions of three variables cannot be represented as compositions of continuous functions of two variables. He obtained the most unexpected result: every continuous function of any number of variables can be represented as a composition of continuous function of three variables. Thus, Hilbert's problem was reduced to a problem of representing functions on universal trees in three-dimensional space. This last problem was solved later, under Kolmogorov's supervision, by his student—Arnold. Finally, Kolmogorov showed that any continuous function can be represented as a composition of continuous functions of a single variable and addition.

Kolmogorov with his students

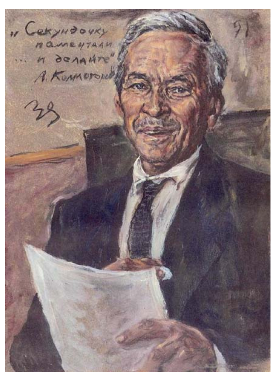
In the 60's Kolmogorov undertook a reconstruction of information theory based on algorithms. He created a new field of mathematics—algorithmic information theory. Kolmogorov's theory states that among all possible algorithmic methods of description, there exist optimal ones with the smallest complexity of its objects.

Mathematical logic (in the broad sense, including the theory of algorithms and the foundations of mathematics) was his first and last love. In 1925, he published a paper on the Law of the Excluded Middle, which has forever become a golden foundation of mathematical logic. This was the first time intuitionistic logic had been systematically researched. With the help of so-called immersion operations (known now as "*Kolmogorov Operations*"), he proved that the application of the Law of the Excluded Middle cannot lead to contradictions. This work, together with the paper published in 1932, made it possible to treat intuitionistic logic as a constructive logic.

A.N. Kolmogorov presenting his lecture

In 1931, Kolmogorov became a *Professor* at Moscow State University, and in 1939, he was elected as *Academician* of the USSR Academy of Sciences. He paid special attention to the training of young scientists and was highly successful in recruiting, among undergraduates and graduates, talented young people fascinated by science. Kolmogorov tried to create a group of research students who would be in a constant state of scientific excitement and continuous research. For all Kolmogorov's students, the years of graduate and post-graduate studies were unforgettable. Their involvement in scientific research was filled with reflections on the role of science. It was also a time of realization and growing faith in the inexhaustible creative power of the human mind.

Above all, Kolmogorov tried to arouse in his students general cultural interests in visual arts, architecture, literature, and even sports. He also created a special high school in Moscow State University: Boarding School 18, or simply "*Kolmogorov School*." The students of this prestigious school systematically took the first places on Russian and International Mathematics and Physics Olympiads. He devoted much of his time to education and improving the teaching of mathematics in the former Soviet Union.



Another painting of Kolmogorov by Dmitrii Gordeev, with the inscription: "*Dream of a second and do it.*"

The list of Kolmogorov's students is extremely large and impressive. Below, we list only those who were elected to different Academies of Sciences:

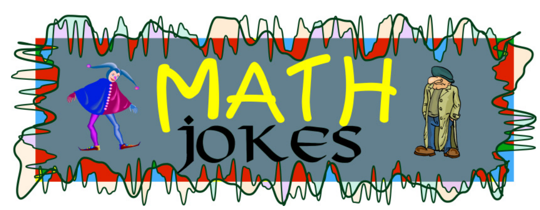
Arnold (Dynamical Systems)
Bol'shev (Mathematical Statistics)
Borovkov (Probability Theory and Mathematical Statistics)
Gelfand (Functional Analysis)
Gnedenko (Probability Theory)
Maltsev (Algebra and Mathematical Logic)
Mikhalevich (Cybernetics)
Millionshchikov (Mechanic and Applied Physics)
Monin (Turbulence and Oceanology)
Nikolskii (Theory of Functions)
Obukhov (Turbulence and Physics of the Atmosphere)
Prokhorov (Probability Theory)
Sevastyanov (Probability Theory)
Shiryaev (Probability Theory and Stochastic Processes)
Sinai (Probability Theory and Dynamical Systems)
Sirachdinov (Probability Theory)

In Moscow State University, Kolmogorov created the Department of Probability Theory, the Department of Mathematical Statistics, and the Department of Mathematical Logic. At the Steklov Mathematical Institute of the Russian Academy of Sciences, he created the Department of Probability Theory and Mathematical Statistics.

His scientific services were highly valued both in his country and abroad. Kolmogorov was awarded many prestigious awards and prizes in the USSR. More than twenty scientific organizations have elected him as a member (Paris Academy of Sciences, London Royal Society, USA National Academy, etc.)

I was very lucky to know A.N. Kolmogorov in person, to whom I was introduced by my supervisor and his former student—A. N. Shiryaev. Later, when I became a member of the Department of Mathematical Statistics at the Steklov Institute of Mathematics, I had an opportunity to work under his direction. As the chair of the Department, he inspired and motivated people to be devoted in their work, for the sake of scientific research. Visiting him occasionally in his apartment in the main building of the Moscow State University, gave me exceptional opportunities to discuss scientific and other topics with this amazing person. We also listened to classical music and read poetry. He loved to work in his country house near Moscow, called "Komarovka," whose ownership he shared with P. S. Alexandrov. During the last years of his life, he spent most of his time in Moscow, unable to visit his much-loved house. Now, due to the efforts of A. N. Shiryaev, this house has been transformed into the Kolmogorov-Alexandrov memorial, which is sometimes open to research visitors from the Steklov Institute of Mathematics and Moscow State University

Andrei Nikolaevich Kolmogorov died on October 20, 1987 and was buried at the Central Moscow Cemetery "*Novodevichye*." The whole life of A.N. Kolmogorov was an unparalleled feat in the name of science.



"Quickie" Inequalities

Murray S. Klamkin†

"Quickie" problems first appeared in the March 1980 issue of Mathematics Magazine. They were originated by the late Charles W. Trigg, a prolific problem proposer and solver who was then the Problem Editor. Many of the first good Quickie proposals were due to the late Leo Moser (who incidentally was a member of the University of Alberta Mathematics Department and subsequently its chairman). These Quickie problems are even now still a popular part of the journal. Also Quickies have proliferated to the problem sections of Crux Mathematicorum, Math Horizons, SIAM Review and Mathematical Intelligencer (unfortunately, no longer in the latter two journals).

Trigg noted that some problems will be solved by laborious methods but with proper insight[1] may be disposed of with dispatch. Hence the name "Quickie".

The probability that two random numbers are equal is zero. It follows that there are more inequalities than equations. Consequently, the study of inequalities are important throughout mathematics. In past issues of *π in the Sky*, December 2001, September 2002, Professor Hrimiuc has provided some good notes on inequalities and we shall be referring to some of them.

Here we illustrate 16 Quickie inequalities and after each one we include for the interested reader an exercise that can be solved in a related manner.

Our first example will set the stage for our Quickie Inequalities.

1. There have been very many derivations published giving the formulas for the distance from a point to a line and a plane. Here is a Quickie derivation for the distance from the point (h, k, l) to the plane $ax + by + cz + d = 0$ in $\mathbb{E}^3$. Here we want to find the minimum value of $[(x-h)^2 + (y-k)^2 + (z-l)^2]^{1/2}$ where (x, y, z) is a point of the given plane. By Cauchy's Inequality,

$$[(x-h)^2 + (y-k)^2 + (z-\ell)^2]^{1/2}[a^2 + b^2 + c^2]^{1/2} \geq |a(x-h) + b(y-k) + c(z-\ell)|$$

or

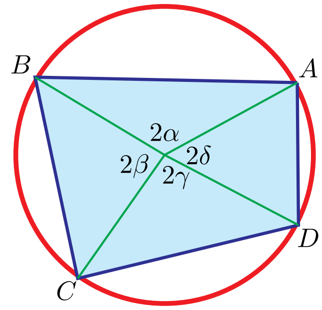
$$\min[(x-h)^2 + (y-k)^2 + (z-\ell)^2]^{1/2} = |ah + bk + c + d|/[a^2 + b^2 + c^2]^{1/2}.$$

Exercise. Determine the distance from the point (h, k) to the line $ax + by + c = 0$.

† Murray S. Klamkin is a professor emeritus at the University of Alberta

[1] and appropriate knowledge-MSK

2. KöMaL problem F. 3097. A convex quadrilateral $ABCD$ is inscribed in a unit circle. Its sides satisfy the inequality $AB \cdot BC \cdot CD \cdot DA \geq 4$. Prove that $ABCD$ is a square.



Let the angles subtended by the four sides from the center be 2α, 2β, 2γ, and 2δ (see figure above). Then $AB = 2\sin\alpha$, $BC = 2\sin\beta$, $CD = 2\sin\gamma$ and $CD = 2\sin\delta$ where $\alpha + \beta + \gamma + \delta = \pi$, $\pi > \alpha, \beta, \gamma, \delta > 0$. Since $\ln(\sin x)$ is concave,

$$\ln(\sin\alpha) + \ln(\sin\beta) + \ln(\sin\gamma) + \ln(\sin\delta) \leq 4\ln\left(\sin\frac{\pi}{4}\right)$$

or $AB \cdot BC \cdot CD \cdot DA \leq 4$. Hence the product is exactly 4 and $\alpha = \beta = \gamma = \delta = \frac{\pi}{4}$ so $ABCD$ is a square.

Exercise. Of all convex n-gons inscribed in a unit circle, determine the maximum of the product of its n sides.

3. KöMaL problem F. 3238. Prove that the inequality

$$\sqrt{a^2 + (1-b)^2} + \sqrt{b^2 + (1-c)^2} + \sqrt{c^2 + (1-a)^2} \geq \frac{3\sqrt{2}}{2}$$

holds for arbitrary real numbers a, b, c.

By Minkowski's Inequality, the sum of the three radicals is grater or equal than $\sqrt{(a+b+c)^2 + (3-a-b-c)^2}$. Then by the power mean inequality or else letting $a + b + c = x$, the expression under the radical is $2(x - 3/2)^2 + 9/2$, so the minimum value is $\frac{3\sqrt{3}}{2}$.

Exercise. Determine the minimum value of

$$\{x^3 + (c-y)^3 + a^3\}^{1/3} + \{y^3 + b^3 + (d-x)^3\}^{1/3},$$

where a, b, c, d are given positive numbers and $x, y \geq 0$.

4. Determine the maximum and minimum z coordinates of the surface

$$5x^2 + 10y^2 + 2z^2 + 10xy - 2yz + 2zx - 8z = 0 \text{ in } \mathbb{E}^3.$$

One method would be to use Lagrange Multipliers. Another more elementary method would be to use discriminants of quadratic equations since if $z = h$ is the maximum, the intersection of the plane $z = h$ with the quadric must be a single point. Even simpler is to express the quadric that is an ellipsoid as a sum of squares, i.e., $(2x+y)^2+(x-y+z)^2+(z-4)^2 = 16$. Hence $\max z = 8$ and $\min z = 0$.

Exercise. Determine the maximum value of y^2 and z^2 where x, y, z are real and satisfy

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

5. Let $a_r = (b_r + b_{r+1} + b_{r+2})/b_{r+1}$ where $b_1, b_2, \ldots, b_n > 0$ and $b_{r+n} = b_r$. Determine the minimum value of

$$\sqrt[3]{a_1} + \sqrt[3]{a_2} + \cdots + \sqrt[3]{a_n}.$$

Even more generally, let

$$x_j = x_{1j} + x_{2j} + \cdots + x_{mj}, \quad j = 1, 2, \cdots, n,$$

where all $x_{ij} > 0$ and $\prod_{j=1}^{n} x_{ij} = P_i^n$, $i = 1, 2, \cdots, m$. Then

$$S \equiv \sqrt[r]{x_1} + \sqrt[r]{x_2} + \cdots + \sqrt[r]{x_n} \ \geq \ n\sqrt[r]{P_1 + P_2 + \cdots + P_m}.$$

We first use the Arithmetic–Geometric Mean Inequality to get

$$S \geq n(x_1 x_2 \cdots x_n)^{1/rn}.$$

Then applying Holder's Inequality we are done. There is equality if and only if $x_{ij} = x_{jk}$ for all i, j, k.

The given inequality corresponds to the special case where $r = m = 3$, $P_1 = P_2 = P_3 = 1$, so that the minimum value is $n\sqrt[3]{3}$.

The inequalities here are extensions of problem #M1277, Kvant, 1991, which was to show that

$$\sum_{i=1}^{n} \{a_i + a_{i+1})/a_{i+2}\}^{1/2} \geq n\sqrt{2}.$$

Exercise. Gy. 2887, KöMal. The positive numbers $a_1, a_2, \ldots, a_n$ add up to 1. Prove the following inequality:

$$(1 + 1/a_1)(1 + 1/a_2) \cdots (1 + 1/a_n) \geq (n + 1)^n.$$

6. If a, b, c are sides of a triangle ABC and R_1, R_2, R_3 are the distances from a point P in plane of ABC to the respective vertices A, B, C. Prove that

$$aR_1^2 + bR_2^2 + cR_3^2 \geq abc.$$

This is a polar moment of inertia inequality and is a special case of the more general inequality

$$(x\vec{\mathbf{A}} + y\vec{\mathbf{B}} + z\vec{\mathbf{C}})^2 \geq 0,$$

where $\vec{\mathbf{A}}$, $\vec{\mathbf{B}}$, $\vec{\mathbf{C}}$ are vectors from P to the respective vertices A, B, C. Expanding out the square, we get

$$x^2R_1^2 + y^2R_2^2 + z^2R_3^2 + 2yz\vec{\mathbf{B}} \cdot \vec{\mathbf{C}} + 2zx\vec{\mathbf{C}} \cdot \vec{\mathbf{A}} + 2xy\vec{\mathbf{A}} \cdot \vec{\mathbf{B}}.$$

Since $2\vec{\mathbf{B}} \cdot \vec{\mathbf{C}} = R_2^2 + R_3^2 - a^2$, etc., the general polar moment of inertia inequality reduces to

$$(x + y + z)(xR_1^2 + yR_2^2 + zR_3^2) \geq yza^2 + zxb^2 + xyc^2.$$

Many triangle inequalities are special cases since x, y, z are arbitrary real numbers. In particular by letting $(x, y, z) = (a, b, c)$, we get our starting inequality. Letting P be the circumcenter and $x = y = z$, we get $9R^2 \geq a^2 + b^2 + c^2$ or equivalently $\sin^2 A + \sin^2 B + \sin^2 C \leq \frac{9}{4}$.

Exercise. Prove that

$$aR_2R_3 + bR_3R_1 + cR_1R_2 \geq abc.$$

7. Prove the identity

$$u(v - w)^5 + u^5(v - w) + v(w - u)^5 + v^5(w - u)$$
$$+ w(u - v)^5 + w^5(u - v) = -10uvw(u - v)(v - w)(w - u),$$

and from this obtain the triangle inequality

$$aR_1(a^4 + R_1^4) + bR_2(b^4 + R_2^4) + cR_3(c^4 + R_3^4) \geq 10abcR_1R_2R_3$$

(with the same notation as in Problem 6).

The identity is a 6th degree polynomial. The left hand side vanishes for $u = 0$, $v = 0$, $w = 0$, $u = v$, $v = w$, and $w = u$. Hence the right hand side equals $kuvw(u - v)(v - w)(w - u)$, where k is a constant. On comparing the coefficients of uv^3w^2 on both sides, $k = -1$.

Now, let u, v, w denote complex numbers representing the vectors from the point P to the respective vertices A, B, C. Taking the absolute values of the both sides of the identity and using the triangle inequality $|z_1 + z_2| \leq |z_1| + |z_2|$, we obtain the desired triangle inequality.

Exercise. Referring to Problem 6, prove that

$$aR_1R_1' + bR_2R_2' + cR_3R_3' \geq abc, \text{ where } R_1', R_2', R_3'$$

are the distances from another point Q to the respective vertices A, B, C.

8. Determine the maximum and minimum values of $x^2 + y^2 + z^2$ subject to the constraint $x^2 + y^2 + z^2 + 2xyz = 1$.

Since it is known that $\cos^2\alpha + \cos^2\beta + \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma = 1$ is a triangle identity, we let $x = \cos\alpha$, $y = \cos\beta$, and $z = \cos\gamma$ where $\alpha + \beta + \gamma = \pi$ and $\pi \geq \alpha$, β, $\gamma \geq 0$. Clearly the maximum of $\cos^2\alpha + \cos^2\beta + \cos^2\gamma$ is 3 and is taken on for $(x, y, z) = (1, 1, -1)$ and permutations thereof. For the minimum (using the above),

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 3 - (\sin^2\alpha + \sin^2\beta + \sin^2\gamma) \geq \frac{3}{4}.$$

Exercise. Determine the maximum of

$$\left\{\sum_{i=1}^{n} x_i\right\}\left\{\sum_{j=1}^{n} \sqrt{a_i^2 - x_i^2}\right\},$$

where $a_i \geq x_i \geq 0$.

9. Problem # 2, Final Round 21st Austrian Mathematical Olympiad. Show that for all natural numbers $n > 2$,

$$\sqrt{2\sqrt[3]{3}\sqrt[4]{4}\cdots\sqrt[n]{n}} < 2.$$

Here we get a better upper bound. If P denotes the left hand side, then

$$\ln P = \frac{\ln 2}{2!} + \frac{\ln 3}{3!} + \cdots + \frac{\ln n}{n!}.$$

Since $\dfrac{\ln x}{x}$ is a decreasing function for $x \geq e$,

$$\ln P < \frac{\ln 2}{2!} + \frac{\ln 3}{3}\left\{\frac{1}{2!}+\frac{1}{3!}+\frac{1}{4!}+\dots\right\}$$
$$= \frac{\ln 2}{2!} + \frac{\ln 3}{3}(e-2) \approx 1.7592.$$

Exercise. Determine a good lower bound for P.

10. Prove that for any distinct real numbers a, b,

$$\frac{e^b - e^a}{b-a} > e^{\frac{b+a}{2}}.$$

This is a special case of the following result due to J. Hadamard [1]: *If a function f is differentiable, and its derivative is an increasing function on a closed interval $[r,s]$, then for all x_1, $x_2 \in (r,s)$ $(x_1 \neq x_2)$, then*

$$\int_{x_1}^{x_2} \frac{f(x)dx}{x_2 - x_1} > f\left\{\frac{x_2+x_1}{2}\right\}.$$

Letting $f(x) = e^x$, we get the desired result.

Exercise. Prove that

$$e^{b^2} - e^{a^2} > (b^2 - a^2)e^{\frac{(b+a)^2}{4}}.$$

11. Prove that

$$\cosh(y-z)+\cosh(z-x)+\cosh(x-y) \geq \cosh x+\cosh y+\cosh z$$

where x, y, z are real numbers whose sum is 0.

Since $\cosh x = \cosh(y+z)$, etc., the inequality can be rewritten as

(i) $\qquad \sinh y \sinh z + \sinh z \sinh x + \sinh x \sinh y \leq 0.$

Since (i) is obviously valid if at least one of $x, y, z = 0$, we can assume that $xyz \neq 0$ and $x, y > 0$. Since $z = -(x+y)$, (i) becomes csch x + csch $y \geq$ csch$(x+y)$ for all $x, y > 0$. This follows immediately since csch t is a decreasing function for all $t > 0$.

Exercise. Prove that

$$\frac{v}{w}+\frac{w}{v}+\frac{w}{u}+\frac{u}{w}+\frac{u}{v}+\frac{v}{u} \geq u+\frac{1}{u}+v+\frac{1}{v}+w+\frac{1}{w}$$

where $u, v, w > 0$ and $uvw = 1$.

12. It is known and elementary that in a triangle, the longest median is the one to the shortest side and the shortest median is the one to the longest side. Determine whether or not the longest median of a tetrahedron is the one to the smallest area face and the shortest median is the one to the largest area face.

Let the sides of tetrahedron $PABC$ be given by $PA = a$, $PB = b$, $PC = c$, $PC = d$, $CA = e$, and $AB = f$. The median m_p from P is given by $\frac{|\vec{\mathbf{A}}+\vec{\mathbf{B}}+\vec{\mathbf{C}}|}{3}$ where $\vec{\mathbf{A}}$, $\vec{\mathbf{B}}$, $\vec{\mathbf{C}}$ are vectors from P to A, B, C respectively. Then

$$9m_p^2 = |\vec{\mathbf{A}}+\vec{\mathbf{B}}+\vec{\mathbf{C}}|^2 = 3(a^2+b^2+c^2) - (d^2+e^2+f^2)$$

and similar formulas for the other medians. It now follows that $9m_a^2 - 9m_b^2 = 4(a^2+f^2) - 4(b^2+e^2)$. It is now possible to have $m_a = m_b$ with their respective face areas unequal, so that the longest median is not one to the smallest face area. The valid analogy is that the longest median is the one to the face for which the sum of the squares of its edges is the smallest, and the shortest median is the one to the face for which the sum of the squares of its edges is the largest.

Exercise. Prove that the four medians of a tetrahedron are possible sides of a quadrilateral.

13. a, b, c, d are positive numbers such that $a^5+b^5+c^5+d^5 = e^5$. Can $a^n+b^n+c^n+d^n = e^n$ for any number $n > 5$?

Let $S_t = x_1^t + x_2^t + \cdots + x_n^t$ where the $x_i \geq 0$. A known result [2] is that the sum $\mathbb{S}_t$ of order t, defined by $\mathbb{S}_t = (S_t)^{1/t}$ decreases steadily from min x_i to 0 as t increases from $-\infty$ to 0^-, and decreases steadily from ∞ to max x_i as t increases from 0^+ to $+\infty$. Consequently, there is no such n.

Exercise. Prove that $\mathbb{S}_T \leq \sum_{i=1}^{n} \alpha_1 \mathbb{S}_{t_i}$ for arbitrary $t_i > 0$ and for $\alpha_1 > 0$, $\sum_{i=1}^{n} \alpha_1 = 1$ and $T = \sum_{i=1}^{n} \alpha_1 t_i$.

14. Prove that

$$\frac{x^{t+1}}{y^t} + \frac{y^{t+1}}{z^t} + \frac{z^{t+1}}{x^t} \geq x+y+z$$

where x, y, $z > 0$ and $t \geq 0$.

Let

$$F(t) = \frac{\left[y\left(\frac{x}{y}\right)^{t+1} + z\left(\frac{y}{z}\right)^{t+1} + x\left(\frac{z}{x}\right)^{t+1}\right]^{\frac{1}{t+1}}}{[x+y+z]^{\frac{1}{t+1}}}.$$

Then by the Power Mean Inequality, $F(t) \geq F(0) = 1$.

Exercise. Prove more generally that

$$\frac{x^{t+1}}{a^t} + \frac{y^{t+1}}{b^t} + \frac{z^{t+1}}{c^t} \geq \frac{(x+y+z)^{t+1}}{(a+b+c)^t},$$

where x, y, z, a, b, $c > 0$ and $t \geq 0$.

15. Determine the maximum value of

$$S = 3(a^3+b^2+c) - 2(bc+ca+ab),$$

where $1 \geq a$, b, $c \geq 0$.

Here, $S \leq 3(a+b+c) - 2(bc+ca+ab)$. Since this latter expression is linear in each of a, b, c, its maximum value is taken on for a, b, $c = 0$ or 1. Hence the maximum is $6-2=4$.

Exercise. Determine the maximum value of

$$S = 4(a^4+b^4+c^4+d^4) - (a^2bc+b^2cd+c^2da+d^2ab)$$
$$-(a^2b+b^2c+c^2d+d^2a),$$

where $1 \geq a, b, c, d \geq 0$.

16. Determine the maximum and minimum values of

$$\sin A + \sin B + \sin C + \sin D + \sin E + \sin F,$$

where $A+B+C+D+E+F=2\pi$ and $\frac{\pi}{2} \geq A, B, C, D, E, F \geq 0$.

Here we get a quick solution by applying Karamata's Inequality [3]. If two vectors $\vec{\mathbf{A}}$ and $\vec{\mathbf{B}}$ having n components, a_i and b_i, are arranged in non-increasing magnitude are such that

$$\sum_{i=1}^{k} a_i \geq \sum_{i=1}^{k} b_i, \quad k=1,2,\ldots,n-1,$$

and

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i,$$

we say that $\vec{\mathbf{A}}$ majorizes $\vec{\mathbf{B}}$ and write $\vec{\mathbf{A}} \succ \vec{\mathbf{B}}$. We then have for a convex function $F(x)$ that

$$F(a_1)+F(a_2)+\cdots+F(a_n) \geq F(b_1)+F(b_2)+\cdots+F(b_n).$$

If $F(x)$ is concave, the inequality is reversed.

Since $\sin x$ is concave in $[0, \pi/2]$, and

$$\left(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}, 0, 0\right) \succ (A,B,C,D,E,F) \succ \left(\frac{2\pi}{6}, \frac{2\pi}{6}, \frac{2\pi}{6}, \frac{2\pi}{6}, \frac{2\pi}{6}, \frac{2\pi}{6}\right).$$

The maximum value is $6\sin\frac{\pi}{3}$ or $\sqrt[3]{3}$ and the minimum value is $4\sin\frac{\pi}{2}$ or 4.

Exercise. Determine the extreme values of $a^5+b^5+c^5+d^5+e^5+f^5$ given that a, b, c, d, e, f, are distinct positive integers with sum 36.

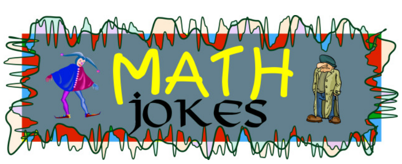
References:

1. D.S. Mitrinovic, *Analytic Inequalities*, Springer-Verlag, Heidelberg, 1970, p. 14.

2. E.F. Beckenbach and R. Bellman, *Inequalities*, Springer-Verlag, New York, 1965, p. 18.

3. A.W. Marshall and I. Olkin, *Inedqualities: Theory of Majorization and Its Applications*, Academic Press, New York, 1979.

Murray S. Klamkin has a long and distinguished career in both industry and academia. He is known primarily as a problem solver, editing the problem corners of many journals over the years. He put this talent to good use in leading the USA team in the IMO, chairing the USAMO committee and authoring several books on mathematics competitions. He is particularly fond of triangle inequalities and spherical geometry. (Andy Liu)



Mother to her daughter: "Why does the tablecloth you just put on the table have the word 'truth' written on it?"

Daughter: "Because I want to turn the table into a truth table!"

Summer Institute for Mathematics at the University of Washington

SIMUW is seeking applications from talented and enthusiastic high school students for its 2004 summer program.

Students experimenting with boomerang.

Admission is competitive. Twenty-four students will be selected from Washington, British Columbia, Oregon, Idaho, and Alaska. Room, board, and participation in all activities are completely free for all admitted participants.

Six weeks of classroom activities, special lectures, and related activities are led by mathematicians and other scientists with the help of graduate and undergraduate teaching assistants.

SIMUW activities are designed to allow students to participate in the experience of mathematical inquiry and to be immersed in the world of mathematics. Topics are accessible yet of sufficient sophistication to be challenging.

2003 SIMUW participants

Students will gain a full appreciation of the nature of mathematics: its wide-ranging content, the intrinsic beauty of its ideas, the nature of mathematical argument and proof, and the surprising power of mathematics within the sciences and beyond.

To obtain more information and application materials, contact us at:

http://www.math.washington.edu/~simuw

SIMUW
Department of Mathematics
University of Washington
Box 354350
Seattle, WA 98195–4350
Phone: (206) 992–5469
Fax: (206) 543–0397
E-mail: simuw@math.washington.edu

The 2004 SIMUW program runs from June 20th to July 31st.

Why I Don't Like "Pure Mathematics"

Volker Runde[†]

I am a pure mathematician, and I enjoy being one. I just don't like the adjective "pure" in "pure mathematics." Is mathematics that has applications somehow "impure"? The English mathematician Godfrey Harold Hardy thought so. In his book *A Mathematician's Apology*, he writes:

> *A science is said to be useful of its development tends to accentuate the existing inequalities in the distribution of wealth, or more directly promotes the destruction of human life.*

His criterion for good mathematics was an entirely aesthetic one:

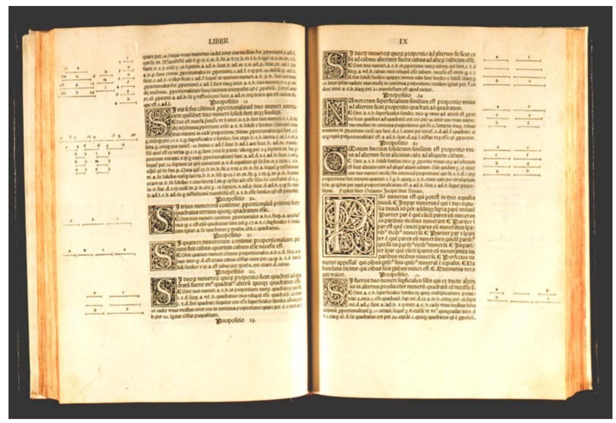
> *The mathematician's patterns, like the painter's or the poet's must be beautiful; the ideas, like the colours or the words must fit together in a harmonious way. Beauty is the first test: there is no permanent place in this world for ugly mathematics.*

I tend to agree with the second quote, but not with the first one.

Godfrey Harold Hardy (1877-1947)

Hardy's book was written in 1940, when the second world war was raging and the memory of the first one was still fresh. The first world war was the first truly modern war in the sense that science was systematically put to use on the battlefield. Physicists and chemists helped to develop weapons of unheard of lethal power. After that war, nobody could claim anymore that science was mainly the noble pursuit of knowledge. Science had an impact on the real world, sometimes a devastating one, and scientist could no longer eschew the moral issues involved. By declaring mathematics—or at least good mathematics—to be without applications, Hardy absolved mathematics, and thus the mathematical community, from being an accomplice of those who waged wars and thrived on social injustice.

The problem with this view is simply that it is not true. Mathematicians live in the real world, and their mathematics interacts with the real world in one way or another. I don't want to say that there is no difference between pure and applied math. Someone who uses mathematics to maximize the time an airline's fleet is actually in the air (thus making money) and not on the ground (thus costing money) is doing applied math, whereas someone who proves theorems on the Hochschild cohomology of Banach algebras (I do that, for instance) is doing pure math. In general, pure mathematics has no *immediate* impact on the real world (and most of it probably never will), but once we omit the adjective *immediate*, the distinction begins to blur.



1533 Edition of Euclid's Elements.

Pierre de Fermat (1601-1665)

The fundamental theorem of arithmetic was already known to the ancient Greeks: every positive integer has a prime factorization that is unique up to the order of the factors. A proof is given in Euclid's more than two thousand years old *Elements*, and there is little doubt that it was known long before it found its way into that book. For centuries, this theorem was the epitomy of beautiful, but otherwise useless mathematics. This changed in the 1970s with the discovery of the RSA algorithm. It is easy to multiply integers on a computer; it is much harder—even though the fundamental theorem says that it can always be done—to determine the prime factorization of a given positive integer. This fact can be used to create codes that are extremely hard to crack. Without them, e-commerce as it exists today would be impossible. Who would want to key his/her credit card number into an online form if he/she had no guarantee that no eavesdropping crook could get hold of it?

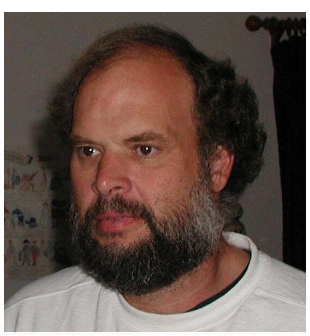
Another mathematical ingredient of the RSA algorithm is Fermat's little theorem (not to be confused with his much more famous last theorem). Pierre de Fermat, a lawyer and

[†] **Volker Runde** is a professor in the Department of Mathematical and Statistical Sciences at the University of Alberta. His web site is http://www.math.ualberta.ca/∼runde/runde.html and his E-mail address is vrunde@ualberta.ca.

civil servant in 17th century France, was doing mathematics in his free time. He did it because he enjoyed the intellectual challenge of it, not because it had any connection with his day job. Here is his little theorem: *If p is a prime number and a is any integer that does not contain p as a prime factor, then p divides $a^{p-1} - 1$.* This theorem is not obvious, but also not very hard to prove (it probably is on the syllabus of every undergraduate course in number theory). Fermat proved it out of curiosity. Computers, let alone e-commerce, didn't exist in his days. Nevertheless, it turned out to be useful more than three hundred years after its creation.

Gottfried Wilhelm von Leibniz (1646-1716)

At the time of Fermat's death, Gottfried Wilhelm von Leibniz was 19 years old. Leibniz would be called, long after his death, the last universal genius: he may have been the last person to have a complete grasp of the amassed knowledge of his time. As a mathematician, he was one of the creators of calculus—no small accomplishment—and he attempted, but ultimately failed, to build a calculating machine, a forerunner of today's computers. As a philosopher, he gained fame (or notoriety) through an essay entitled *Théodizée* (meaning God's defense) in which he tried to reconcile the belief in a loving, almighty God with the apparent existence of human suffering: he argued that we do indeed live in the best of all possible worlds. Philosophical and theological considerations led him to discover the binary representation of numbers: instead of expressing a number in the decimal system, e.g., $113 = 1 \cdot 10^2 + 1 \cdot 10 + 3 \cdot 10^0$, we can do it equally well in the binary system ($113 = 1 \cdot 2^6 + 1 \cdot 2^5 + 1 \cdot 2^4 + 0 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2 + 1 \cdot 2^0$). Since numbers in binary representation are easy to implement on electronic computers, Leibniz's discovery helped to at least facilitate the inception of modern information technology.



Vaughan Jones

Almost three hundred years after Leibniz's death, the mathematician Vaughan Jones was working on the problem of classifying subfactors (I won't attempt to explain what a subfactor is; it has nothing to do with multiplying numbers). To accomplish this classification, he introduced what is now called the Jones index: with each subfactor is associated a certain number. This index displays a rather strange behaviour, it can be infinity or any real number greater than or equal to 4, but the values it can attain under 4 have to be of the form $4\cos^2(\pi/n)$ with $n = 3, 4, \ldots$. Jones asked himself why. His research led to the discovery of the Jones polynomial (of course, he didn't call it that) for which he was awarded the Fields Medal, the highest honour that can be bestowed upon a (pure) mathematician. This Jones polynomial, in turn, has helped molecular biologists to better understand the ways DNA curls up in a cell's nucleus.

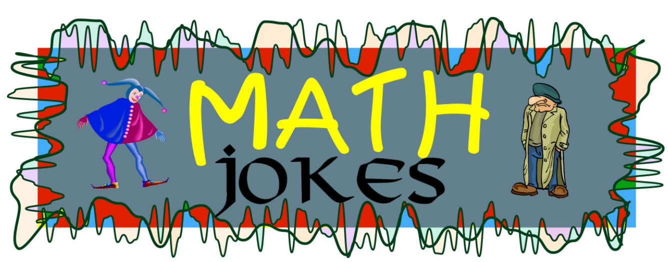
Most of pure mathematics will probably never impact the world outside the mathematical community, but who can be sure in a particular case? In the last twenty five years, the intellectual climate in most "developed" countries has become increasingly unfavourable towards *l'art pour l'art*. Granting agencies nowadays demand that researchers explain what the benefits of their research are to society. In principle, there is nothing wrong with that; taxpayers have a right to know what their money is used for. The problem is the time frame. The four examples I gave portray research that was done for nothing but curiosity and the sheer pleasure of exploration, but that turned out nonetheless to have applications with sometimes far reaching consequences. To abandon theoretical research just because it doesn't have any foreseeable application in the near future is a case of cutting off one's nose despite the face.

Pure mathematics isn't pure: neither in the sense that it is removed from the real world, nor in the sense that its practitioners can ultimately avoid the moral questions faced by more applied scientists. A more fitting title might be"theoretical mathematics."

Enigma Machine used by German Navy

P.S. While Hardy wrote his *Apology*, other British mathematicians worked on and eventually succeeded in breaking the Enigma Code used by the German navy. By all likelihood, their work helped shorten the war by months if not years, thus saving millions of lives on both sides.

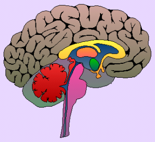
P.P.S. In 1908, Hardy came up with a law that described how the proportions of dominant and recessive genetic traits are propagated in large populations. He didn't think much of it, but it has turned out to be of major importance in the study of blood group distributions.



Mathematics is made of 40 percent formulas, 40 percent proofs and 40 percent imagination.

Q: What caused the big bang?
A: God divided by zero. Oops!

Math is like love: a simple idea but it can get complicated.



Math Challenges

Problem 1. Find all functions $f : (0, \infty) \to (0, \infty)$ such that

$$f\big(x + \sqrt{x}\big) \le x \le f(x) + \sqrt{f(x)} \quad \text{for all } \; x \in (0, \infty).$$

Problem 2. Find all distinct pairs (x, y) of integers that are solutions of the equation

$$x^2 - xy + y^2 = x + y.$$

Problem 3. Find the largest subset $A \subset \{1, 2, \ldots, 2003\}$ such that for all a, $b \in A$, $a + b$ is not divisible by $a - b$.

Problem 4. Let x_1, x_2, $\ldots$, x_{2004} be positive real numbers such that

$$\frac{1}{2003 + x_1} + \frac{1}{2003 + x_2} + \cdots + \frac{1}{2003 + x_{2004}} > 1.$$

Prove that $x_1 x_2 \cdots x_{2004} < 1$.

Problem 5. Four points are given in the plane. If the distance between any two of them is not less then $\sqrt{2}$ and not greater than 2, prove that these points are the vertices of a square.

Problem 6. Find the maximum value of the area of a triangle ABC that has vertices on three circles centered at the same point with radii 1, $\sqrt{7}$, and 4, respectively.

Send your solutions to π *in the Sky*: Math Challenges.

Solutions to the Problems Published in the September, 2002 Issue of π *in the Sky*:

Problem 1. Let n be a fixed positive integer and consider the more general problem of solving

$$\frac{xy}{x + y} = n,$$

where x and y are positive integers. Then $y = \frac{nx}{x-n}$. In particular, y is a positive integer if and only if $x - n$ is a positive integer that divides nx. But $nx = n^2 + n(x - n)$, so we see that $x - n$ divides nx if and only if it divides n^2. Consequently, the number of x values yielding positive integers y is precisely equal to the number of positive divisors of n^2. Indeed, for each positive divisor d of n^2, we let $x = n + d$. For example, when $n = 100$, we get

$$n^2 = 100^2 = 2^4 \cdot 5^4.$$

Thus, the positive divisors of 100^2 are precisely the numbers of the form $2^a \cdot 5^b$ with $a = 0, 1, 2, 3$, or 4, and $b = 0, 1, 2, 3$, or 4. It follows that there are $5 \cdot 5 = 25$ divisors and hence there are 25 positive integers x that yield positive integer y.

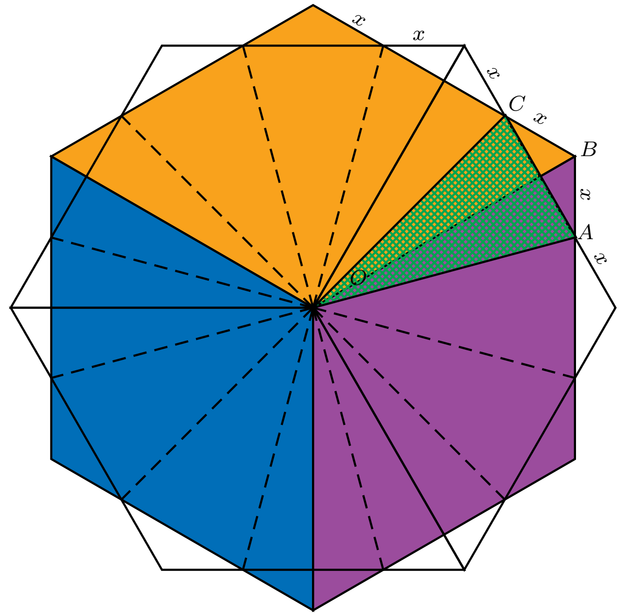
Problem 2. If $2003 + n = m(n + 1)$, then $2002 = m(n + 1) - (n + 1) = (m - 1)(n + 1)$ and $n + 1 \ge 1$ is a divisor of $2002 = 2 \cdot 7 \cdot 11 \cdot 13$. In particular,

$$n + 1 = 2^a \cdot 7^b \cdot 11^c \cdot 13^d$$

with each a, b, c, d being 0 or 1. Since there are $2^4 = 16$ possible choices for the exponent, there are 16 possible choices for n.

The above solutions of the problems 1 and 2 were presented to π *in the Sky* by Jeganathan Sriskandarajah from Madison, WI. These problems were also correctly solved by Robert Bilinski from Montréal and Edward T.H. Wang from Waterloo.

Problem 3. (**Solution by Wieslaw Krawcewicz**) The picture below illustrates the solid P that is the intersection of the unit cube with a copy that is rotated 30 degrees. This solid can be obtained by cutting off from the initial cube six identical tetrahedrons, one of which, denoted by $OABC$, is indicated in the picture.



Th triangle ABC is right-angled and it has sides x, x, and $\sqrt{2}x$. Since the length of the edge of the cube is one, we get

$$1 = x + \sqrt{2}x + x \iff x = \frac{1}{2 + \sqrt{2}}.$$

Therefore the volume of the tetrahedron $OABC$ is $\frac{1}{6}x^2$, and consequently we find that the volume V of the solid P is given by

$$V = 1 - 6\left(\frac{1}{6} \cdot \frac{1}{(2 + \sqrt{2})^2}\right) = \frac{5 + 4\sqrt{2}}{6 + 4\sqrt{2}}.$$

Madison Area Technical College Math Club Celebrated π Day on March 14, 2003

The main event of the π Day celebration was the math competition featuring teams from six different two-year colleges in Wisconsin. Among the other activities, there was also an informative presentation on "The Calculation of Pi," a poster competition and a pie-eating contest. At an awards ceremony the winning teams and individuals were presented with their prizes. The picture above shows several contestants at the beginning of the math competition.