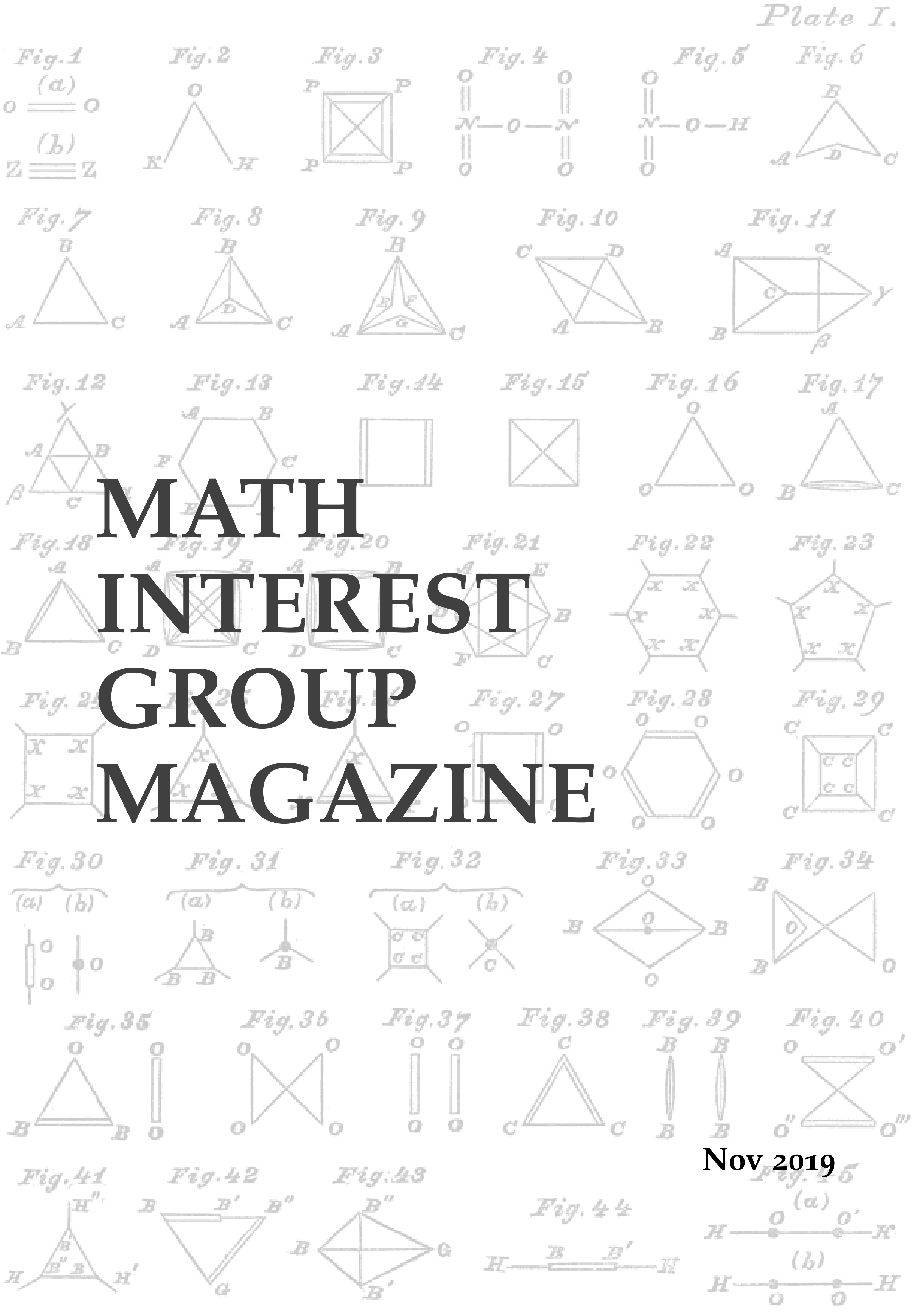


MATH INTEREST GROUP MAGAZINE

Nov 2019

Preface

Ladies and gentlemen, we are back with another issue of the MIG magazine. Apologies for the delay; I know you've been spending sleepless nights waiting for this masterpiece.

Compared to the other interest groups, it is my opinion that MIG lacks presence in our school. That's why we're spicing things up. Maybe one day we'll get funding for pizza, like some other groups. . .

One last thing—we've made some decisions regarding the target audience of the magazine. Originally, we intended for every article to be readable by as many people as possible, but this was unrealistic and restrictive as it prevented us from writing on more interesting topics. As such, you'll find that some articles here are harder to read than others, and possibly out of your grasp. This is perhaps for two reasons that aren't necessarily mutually exclusive—the inadequacies of the author, or the inadequacies of the reader. But I digress. On with the show!

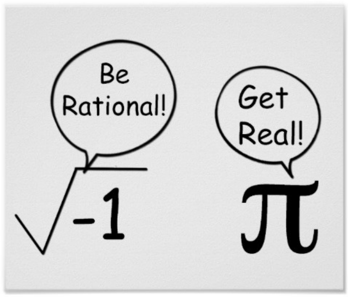
— The Editor

Cover illustration is from J. J. Sylvester (1878). *On an Application of the New Atomic Theory to the Graphical Representation of the Invariants and Covariants of Binary Quantics, with Three Appendices.* American Journal of Mathematics **1**(1), pp. 64–104. In this paper, Sylvester coins the term *graph* as in the study of graph theory, which seeks to understand networks and their connections mathematically.

π

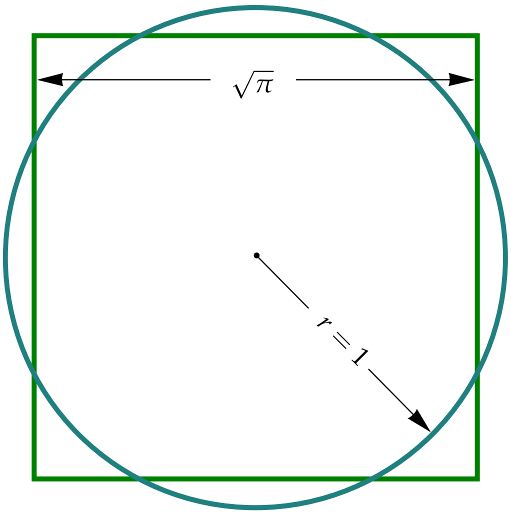
Gnoh Cheng Yi

Looking at the school building from the field, what do you see? The "greenush" pi wall! For those who don't know, the pi wall represents $\pi = 3.14159\ldots$, with consecutive tiles being of the same color. The number of consecutive tiles represents a digit of π. For example, the first few digits of π are represented as follows:



π is a mathematical constant originally defined as the ratio of a circle's circumference to its diameter. It has several properties that make it very different from a "normal" number. Most of you probably know that it is irrational, not because it can't think rationally, but because it cannot be expressed as a ratio of two integers.[1]

But do you know that π is also transcendental?



Given a circle, it is impossible to use a compass and straightedge (a ruler with no markings) to construct a square with the same area as the circle.

π being transcendental means that it is impossible to find any nonzero polynomial equation with integer coefficients such that π is a root. Another well-known transcendental number is e, Euler's number. A consequence of π being transcendental is that it is impossible to construct, using compass and straightedge alone, a square whose area is exactly equal to the area of a given circle.

[1] For those interested, some proofs of the irrationality of π can be found at https://en.wikipedia.org/wiki/Proof_that_pi_is_irrational.

Another interesting property of π is that it can be expressed as an infinite series of nested fractions, called a continued fraction:

$$\pi = 3 + \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{1 + \cfrac{1}{292 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \ddots}}}}}}}$$

Approximations of π can be obtained by truncating the continued fraction at any point. Some of the approximations include 3, $22/7$, $333/106$, and $355/113$.

There are several world records associated with π. One of the records, on the approximation of π,[2] was recently created in March 2019. The record of $31,415,926,535,897 = \lfloor \pi \times 10^{13} \rfloor$ digits was attained by Google employee Emma Haruka Iwao using the Google Cloud Compute Engine over the span of 121 days.

[2] https://en.wikipedia.org/wiki/Chronology_of_computation_of_pi

Another world record is the most digits of π memorised.[3] Guinness World Records recorded a stunning record of 70,000 digits by Rajveer Meena from India when most of us can't even memorise 70 digits of π. The whole challenge took nearly 10 hours.

[3] https://www.guinnessworldrecords.com/world-records/most-pi-places-memorised

Here are a few challenges for you:

1. Find the first decimal place where the digit 0 appears in π.

2. Memorise 70 digits of π. ("NUS High, NUS High, we know all the digits of π . . . ")

Let's derive some infinite series!

Raghavendra Narayan Rao

Does this ring a bell? Consider $\sum_{k=1}^{\infty} k \overset{??}{=} -1/12$. Top 10 anime betrayals indeed, but that's not what we are here for. Let's look at some "real" and convergent infinite series.

Ooh, this boy looks like a good candidate: $\sum_{k=1}^{\infty} 1/k$.

Wait... that's divergent, unfortunately. We have

$$\begin{aligned} & 1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}+\dots \\ > & 1+\frac{1}{2}+\frac{1}{4}+\frac{1}{4}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\dots \\ = & 1+\frac{1}{2}+\frac{1}{2} \qquad +\frac{1}{2} \qquad\qquad\qquad +\dots, \end{aligned}$$

which grows to infinity; we say that it diverges.

This just showed that infinite series can diverge although terms can get arbitrarily close to 0.

What about $\sum_{k=1}^{\infty} \frac{1}{k^2} = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \dots$? Turns out that it is in fact convergent, and further, equal to $\pi^2/6$. Wow.

Here's another two infinite series[4] that fascinate me:

$$\begin{cases} 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \frac{\pi}{4} \\ 1 - \frac{1}{3^2} + \frac{1}{5^2} - \frac{1}{7^2} + \cdots = \frac{\pi^2}{8} \end{cases}$$

These wacky series boys all have one thing in common: π appears, somehow! It turns out that we can prove these results using something called a *Fourier series*. What's more, you'll only formally need year 3 mathematics to understand this[5] and go on to evaluate other fantastic infinite series!

To begin, let's define a function $f : \mathbf{R} \to \mathbf{R}$ as such:

$$\begin{aligned} f(x) = {} & a_0 \cos 0x + a_1 \cos 1x + a_2 \cos 2x + \cdots + a_n \cos nx + \cdots \\ & + b_0 \sin 0x + b_1 \sin 1x + b_2 \sin 2x + \cdots + b_n \sin nx + \cdots, \end{aligned}$$

where $a_0, a_1 \dots$ and $b_0, b_1, \dots$ are constants to be determined. Now, this simplifies to

$$\begin{aligned} f(x) = {} & a_0 + a_1 \cos x + \cdots + a_n \cos nx + \cdots \\ & + b_1 \sin x + b_2 \sin 2x + \cdots + b_n \sin nx + \cdots. \end{aligned}$$

This function can represent *any*[6] periodic function, depending only on how you tweak the constants. For convenience, we shall deal

[4] **Joke 1.1.** Is the plural form 'serieses'? Or 'serii'?
Well, who knows? Certainly not I, Raghavendra Narayan Rao. This is the *math* interest group. □

[5] Or at least, trick yourself into believing that you understand.

[6] The Editor: There *are* technical details, but any nice function (say, twice continuously differentiable) that you can think of will do just fine. See https://en.wikipedia.org/wiki/Convergence_of_Fourier_series.

with functions of period 2π. That is, functions like sin and cos which repeat themselves every 2π.[7]

Example. The *square wave* is this friendly boy over here.

Before we go further, here is some advice—you should be familiar with your factor formulas, as well as how $\int_0^{2\pi} \cos mx\, dx = \int_0^{2\pi} \sin mx\, dx = 0$ for m a non-zero integer.[8] Hopefully $\int_0^{2\pi} a\, dx = 2\pi a$ is clear to you.[9] Also,

$$\begin{aligned}\int_0^{2\pi} \cos mx \cos nx\, dx &= \frac{1}{2}\int_0^{2\pi} \cos(m+n)x + \cos(m-n)x\, dx \\ &= \frac{1}{2}\int_0^{2\pi} \cos(m+n)x\, dx + \frac{1}{2}\int_0^{2\pi} \cos(m-n)x\, dx.\end{aligned}$$

We're almost done. About three more things you should know. If $m \neq n$, then

$$\frac{1}{2}\int_0^{2\pi} \cos(m+n)x\, dx + \frac{1}{2}\int_0^{2\pi} \cos(m-n)x\, dx = 0.$$

If $m = n$, then

$$\begin{aligned}\frac{1}{2}\int_0^{2\pi} \cos(m+n)x\, dx + \frac{1}{2}\int_0^{2\pi} \cos(m-n)x\, dx &= \frac{1}{2}\int_0^{2\pi} \cos(m-n)x\, dx \\ &= \frac{1}{2}(2\pi) = \pi.\end{aligned}$$

Similarly, applying factor formulas gives us

$$\int_0^{2\pi} \sin mx \sin nx\, dx = \begin{cases} 0, & m \neq n \\ \pi, & m = n \end{cases}$$

and

$$\int_0^{2\pi} \sin mx \cos nx\, dx = 0$$

for all $m, n \in \mathbf{Z}$.

There. Now that we've lost 99% of our audience, it's time for us to lose the rest of it. Indeed, integrating $f(x)$ from 0 to 2π, we find that

$$\begin{aligned}\int_0^{2\pi} f(x)\, dx = \int_0^{2\pi} a_0\, dx \quad &+ \int_0^{2\pi} a_1 \cos x\, dx + \cdots + \int_0^{2\pi} a_n \cos nx\, dx + \cdots \\ &+ \int_0^{2\pi} b_1 \sin x\, dx + \cdots + \int_0^{2\pi} b_n \sin nx\, dx + \cdots \\ = 2\pi a_0.\end{aligned}$$

[7] More formally, a function $f : \mathbf{R} \to \mathbf{R}$ is said to be 2π-periodic if $f(x + 2\pi) = f(x)$ for every $x \in \mathbf{R}$.

[8] From here on out we will always use m and n as non-negative integers unless otherwise stated.

[9] If not, please consult your math teacher!

Further, we have

$$\begin{aligned}\int_0^{2\pi} f(x)\cos mx\,dx &= \int_0^{2\pi} a_0\cos mx\,dx + \int_0^{2\pi} a_1\cos x\cos mx\,dx + \cdots \\ &+ \int_0^{2\pi} a_m\cos mx\cos mx\,dx + \cdots \\ &+ \int_0^{2\pi} b_1\sin x\cos mx\,dx + \cdots + \int_0^{2\pi} b_m\sin mx\cos mx\,dx + \cdots \\ &= \int_0^{2\pi} a_m\cos mx\cos mx\,dx \\ &= \pi a_m\end{aligned}$$

and

$$\begin{aligned}\int_0^{2\pi} f(x)\sin mx\,dx &= \int_0^{2\pi} a_0\sin mx\,dx + \int_0^{2\pi} a_1\cos x\sin mx\,dx + \cdots \\ &+ \int_0^{2\pi} a_m\cos mx\sin mx\,dx + \cdots \\ &+ \int_0^{2\pi} b_1\sin x\sin mx\,dx + \cdots + \int_0^{2\pi} b_m\sin mx\sin mx\,dx + \cdots \\ &= \int_0^{2\pi} b_m\sin mx\sin mx\,dx \\ &= \pi b_m.\end{aligned}$$

Now that I've shown you some equations without giving you any intuition whatsoever, it follows that

$$a_0 = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,dx,$$

$$a_m = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos mx\,dx \quad\text{and}\quad b_m = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin mx\,dx.$$

Let's look at an example.

Example. Let $f : \mathbf{R} \to \mathbf{R}$ be defined on $[0, 2\pi]$ by

$$f(x) = \begin{cases} 1, & 0 \le x < \pi \\ 0, & \pi \le x < 2\pi \end{cases}$$

and extended 2π-periodically to the rest of $\mathbf{R}$; i.e., by requiring that $f(x + 2\pi) = f(x)$ for all $x \in \mathbf{R}$. By evaluating the Fourier series of f, prove that

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots.$$

Solution. We have

$$\begin{aligned}a_0 &= \frac{1}{2\pi}\int_0^{2\pi} f(x)\,dx \\ &= \frac{1}{2\pi}\int_0^{\pi} 1\,dx + \frac{1}{2\pi}\int_\pi^{2\pi} 0\,dx \\ &= \frac{1}{2}.\end{aligned}$$

Also,

$$\begin{aligned} a_m &= \frac{1}{\pi}\int_0^{2\pi} f(x)\cos mx\,dx \\ &= \frac{1}{\pi}\int_0^{\pi} \cos mx\,dx + \frac{1}{\pi}\int_{\pi}^{2\pi} 0\,dx \\ &= \frac{1}{\pi}\left[\frac{\sin mx}{m}\right]_0^{\pi} \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} b_m &= \frac{1}{\pi}\int_0^{2\pi} f(x)\sin mx\,dx \\ &= \frac{1}{\pi}\int_0^{\pi} \sin mx\,dx + \frac{1}{\pi}\int_{\pi}^{2\pi} 0\,dx \\ &= \frac{1}{\pi}\left[\frac{-\cos mx}{m}\right]_0^{\pi} \\ &= \frac{1}{m\pi}(\cos 0 - \cos m\pi) \\ &= \begin{cases} 0, & m \text{ is even,} \\ \frac{2}{m\pi}, & m \text{ is odd.} \end{cases} \end{aligned}$$

It follows that

$$f(x) = \frac{1}{2} + \frac{2}{\pi}\left(\frac{\sin x}{1} + 0 + \frac{\sin 3x}{3} + 0 + \frac{\sin 5x}{5} + 0 + \cdots\right).$$

In particular, evaluating f at $\pi/2$ gives us

$$1 = \frac{1}{2} + \frac{2}{\pi}\left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots\right),$$

and so the result follows. □

Here are some results you may try to derive using the ideas developed above.

Exercise. Let $f : \mathbf{R} \to \mathbf{R}$ be defined by

$$f(x) := \begin{cases} x, & 0 \le x < \pi \\ \pi, & \pi \le x < 2\pi, \end{cases}$$

on $[0, 2\pi]$ and extended 2π-periodically to the rest of $\mathbf{R}$. By evaluating the Fourier series of f, prove that

$$1 - \frac{1}{3^2} + \frac{1}{5^2} - \frac{1}{7^2} + \cdots = \frac{\pi^2}{8}.$$

Exercise. Let f be the function defined on $[-\pi, \pi]$ by $f(\theta) = |\theta|$. Evaluate the Fourier series of f at $\theta = 0$ and hence show that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

My name is Raghavendra Narayan Rao, and thank you for coming to my TED talk.[10]

[10] THE EDITOR: C, D, E♭, F, D, B♭, C

Primality Testing

Ashley Aragorn Khoo

A primality test is an algorithm for determining whether a given number is prime or not. Firstly, this is unlike integer factorization in which the prime factors of a given number need to be found. Whereas for primality tests, it is only required for one to figure out whether a number is prime or not. Another difference between primality tests and integer factorization is that integer factorization is probably computationally difficult while primality testing is "fast" (It can be done in polynomial time).

Some readers may be aware of an algorithm known as the sieve of Eratosthenes which can quickly find all prime numbers between 1 and n. The first thing to do is to list all the integers between 2 to n. Then, we go through the numbers we have written down from smallest to largest. If the number has not been crossed out, it is a prime. We will then cross out all multiples of that number we have declared as prime. The proof of this algorithm is left to the reader.

Can we do better? Yes.

Firstly, we do not need to check if all the numbers between 1 to n are prime just to find out if n is prime. So, we shall use the definition of a prime number. According to WolframAlpha[11], "An integer greater than 1 is prime if its only positive integer divisors are 1 and itself. Otherwise, it is composite." So we can just check for all integers in the range $[2, n)$ whether they are divisors of n to find out if n is a prime number.

Can we do better? Yes.

Well, if we think about it hard enough we realize that if a number is composite, then surely its smallest factor greater than 1 is at most its square root. So now we only need to check in the range of $[2, \sqrt{n}]$. This is implemented below in `isPrime_naive()`.

Can we do better? Yes.

Unfortunately, the next method discussed is probabilistic. Which means that there is a chance that the method is wrong, but the chance is small in practice, so we do not care, or a counterexample has not been found. Although, for those interested there is a paper on a fast and deterministic primality test.[12]

The probabilistic algorithm to be discussed will be the Miller–Rabin test, which relies on Fermat's little theorem, which states that $\alpha^p \equiv \alpha \pmod{p}$, if α and p are coprime.[13]

Using Fermat's little theorem, we get a result that if $\alpha^{p-1} \not\equiv 1 \pmod{p}$, then α and p are definitely not coprime. However it should

[11] https://www.wolframalpha.com/examples/mathematics/number-theory/prime-numbers/

[12] Agrawal, M., Kayal, N. and Saxena, N. (2004). *PRIMES is in* P. Annals of Mathematics **160**, pp. 781–793.
Available online at http://annals.math.princeton.edu/2004/160-2/p12.

[13] There is a beautiful way of thinking about Fermat's little theorems where we can visualize it with coloured beads. See Brit Cruise's video https://www.youtube.com/watch?v=XPMzosLWGHo.

be noted that if $\alpha^{p-1} \equiv 1 \pmod p$, then α and p may not be coprime.

So now, we get into the essence of the Miller–Rabin test, in which for some number p, we randomly select an integer, α, in the range $[2, p-1]$. Then, we compute α^{p-1}. If the result is not $1 \pmod p$, the number is composite. Else, the number is probably prime. Furthermore, we can do an improvement, where if we represent p as $2^s d + 1$ where d is an odd number, for $0 \leq k < s$ where if $\alpha^{2^k d} \not\equiv \pm 1 \pmod p$ and $\alpha^{2^{k+1} d} \equiv 1 \pmod p$, then p is not a prime number. The proof of this is again left as an exercise to the reader. If we have determined that p is composite using some α, we denote $f(\alpha, p) = 1$.

Now, we just keep guessing for α in the above range. It can be shown for some composite p that at least $3/4$ of the integers, $f(\alpha, p) = 1$.[14] So if we run this algorithm k times, we are sure that we are with a probability of at least $1 - (1/4)^k$ that our result is correct. It is also notable that if we test for all integers, α, in the range $[2, 2(\ln p)^2]$, there will exist an α such that $f(\alpha, p) = 1$ holds for all composite p if the Generalized Riemann Hypothesis[15] is true.[16] However for smaller numbers, the number of integers, α, one has to check for is significantly reduced. In fact, only checking the integers 2, 3, 5 and 7, one can accurately check for prime numbers up to $3, 215, 031, 751$.[17] This is implemented below in `isPrime_fast()`. Can we do better? Yes.

If a reader has access to Mathematica, I recommend them trying to put a 2000 digit number into the `PrimeQ()` function which is the primality test function in Mathematica. It is amazing how fast it can deduce whether the input is prime or not. After referring to Mathematica's implementation[18], it performs a Miller–Rabin test with α being 2 and 3, then performing a Lucas test.[19] This is very fast and some friends[20] and I were able to generate a 2005-digit prime number which we call "The NUSH Prime" which is on the next page.

Can we do better? Maybe; who knows?

[14] Rabin, M. O. (1980). *Probabilistic algorithm for testing primality*. Journal of Number Theory **12**(1), pp. 128–138. Available online at https://www.sciencedirect.com/science/article/pii/0022314X80900840.

[15] http://mathworld.wolfram.com/GeneralizedRiemannHypothesis.html

[16] Bach, E. (1990). *Explicit bounds for primality testing and related problems*. Math. Comp. **55**, pp. 355–380. Available online at https://www.ams.org/journals/mcom/1990-55-191/S0025-5718-1990-1023756-8/.

[17] Pomerance, C., Selfridge, J. L. and Wagstaff, S. S. (1980). *The pseudoprimes to* $25 \cdot 10^9$. Math. Comp. **35**, pp. 1003–1026. Available online at https://www.ams.org/journals/mcom/1980-35-151/S0025-5718-1980-0572872-7/.

[18] https://reference.wolfram.com/language/tutorial/SomeNotesOnInternalImplementation.html

[19] https://en.wikipedia.org/wiki/Lucas_primality_test

[20] Lin Yicheng, Lim Huai Kai, Lam Jun Rong, Daniel Ng Chun Kit & Dylon Wong Yee Kin

Here is a fast primality testing script written in `Python`.

```python
def isPrime_naive(p):
    if (p==1):
        return False
    if (p==2):
        return True
    if (p%2==0):
        return False
    for i in range(3,int(p**0.5+1),2):
        if (p%i==0):
            return False
    return True



def exp(base, power, mod): #quickly calculates (base^power)%mod
    if (power==1): #base case
        return base
    elif(power%2==0): #if power is even, (base^power)=(base^(power/2))^2
        return (exp(base,power//2,mod)**2)%mod
    else:
        return ((exp(base,power//2,mod)**2)*base)%mod

bases=[2,3,5,7] #this should only contain prime numbers
#for testing bases 2,3,5,7 it is verified to work until 3215031751
def isPrime_fast(p):
    if (p==1):
        return False
    for i in range(0,len(bases)):
        if (p==bases[i]):
            return True
    if (p%2==0):
        return False
    d=p//2
    s=1
    while (d%2==0):
        d//=2
        s+=1
    for i in range(0,len(bases)):
        res=exp(bases[i],d,p)
        for j in range(s):
            res2=(res*res)%p
            if (res2==1 and (res!=1 and res!=p-1)):
                return False
            res=res2
        if (res!=1):
            return False
    return True
```

```
1111111111111111111111111111111111
1111111111111111111111111111111111
1111111111111111111111111111111111
1188881111111111111111118811111111
1188881111111111111111188881111111
1188881111111111118111188881111811
1188881111111111188811188881118881
1188881111111111188888188881888881
1188881111111111188888888888888881
1188881111111111118888888888888811
1188881111111111111188888888881111
1188881111111111111188888888881111
1188881111111111118888888888888811
1188881111111111188888888888888881
1188881111111111188888188881888881
1188881111111111188811188881118881
1188881111111111118111188881111811
1188881111111111111111188881111111
1188881111111111111111118811111111
1188881111111111111111111111111111
1188881111111111111111111111111111
1188881111111111111111111111111111
1188881111188888888888888888888111
1188881111188888888888888888888111
1188881111188888888888888888888111
1188881111111118888111188881111111
1188881111111118888111188881111111
1188881111111118888111188881111111
1188881111111118888111188881111111
1188881111111118888111188881111111
1188881111111118888111188881111111
1188881111111111111111111111111111
1188881111111111111111188881111111
1188881111111111111111188881111111
1188881111111111111111188881111111
1118888111111111111111888811111111
1118888111111111111111888811111111
1111888811111111111118888111111111
1111888811111111111118888111111111
1111188881111111111188881111111111
1111118888888888888888811111111111
1111111888888888888888111111111111
1111111111888888888111111111111111
1111111111111111111111111111111111
1111111111111111111111111111111111
0000000000000000000000000000000000
0000000000000000000000000000000000
0000000000000000000000000000000000
0000000000000000000000000000000000
0000000000000000000000000000000000
0000000000000000000000000000000000
0000000000000000000000000000000000
0000000000000000000000000000000000
0000000000000000000000000000000000
0000000000000000000000000000000000
0000000000000000000000000000000000
0000000000000000000000000000000000
0000000000000000000000000000000000
000000000000000000000000000002101
```

All 2005 digits of the NUSH prime. An online copy can be found at `https://tinyurl.com/nushprime`.

Probability is confusing

Paul Seow Jian Hao

Probability problems can often be confusing, so let's introduce it insufficiently, then throw you some puzzles.

For the sake of first-timers, what do we mean when we talk about probability? In mathematics, it refers to quantifying the likelihood of a certain event occuring. Often, we let this event A have a probability $P(A)$, and we assign this a value between 0 and 1, with 0 meaning it has no chance of occuring, and 1 representing it definitely occuring.[21] For example, if a meteorologist says "there is a 60% chance of rain today", if we let the event A be "it will rain today", then $P(A) = 0.6$.

[21] The Editor: Okay... Technically this isn't right. Probability is actually way more confusing than this, and the main idea is roughly that points have no area, but modern mathematics treats areas as collections of points, so this leads to something that feels paradoxical, much like the age-old question of how many grains of sand are needed to form a pile of sand—three grains surely isn't a pile. What about four? For more details, a keyword to look out for is "measure-theoretic probability'.

In many simple problems, we can also calculate a probability by finding the number of desired outcomes and dividing by the total number of outcomes. We can also find the complement of A, given by $P(\text{not } A) = 1 - P(A)$, which is quite an intuitive definition.

Often, we are more interested in the probability of multiple events happening together, rather than just one. There's a few things we should take note of before calculating our probability:

1. **Mutually exclusive events.** This refers to two events that could occur separately but not simultaneously in a single experiment. For example, if we have a dice roll, getting a 1 and 6 at the same time is impossible, so they are mutually exclusive.

   Hence, given two mutually exclusive events A and B, we have
   $$P(A \text{ or } B) = P(A) + P(B),$$
   since $P(A \text{ and } B) = 0$.

2. **Not mutually exclusive events.** If two events aren't mutually exclusive, then we can write
   $$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B).$$
   This is known as the principle of inclusion and exclusion, and can be visualised as an "overlap" between the two overlaps being removed. We can also now clearly see why we can just add $P(A)$ and $P(B)$ for mutually exclusive events.

3. **Independent events.** Two events are independent if they have no relation or connection to each other. For example, the author getting a 5.0 for HCL (with probability 0) and the author writing this article on time (hopefully with probability 1) are independent. We can then calculate the joint probability by
   $$P(A \text{ and } B) = P(A)P(B).$$

Simple? Well here's a couple of easy problems. (The solutions are at the end of this article.)

Problem 1. Suppose I have a bag of 2 red balls, 3 blue balls and 4 green balls. Given that I pick 3 balls at random, what's the probability I pick out 1 ball of each colour?

Problem 2. Find the probability that a leap year has 52 Sundays.

Too simple? Well, here's a slightly more complicated problem:

Problem 3. A bag contains one marble, either black or white, with equal probability. A new black marble is now added to the bag, which is then shaken. When a marble is taken out at random, it turns out to be black. What is the probability that the remaining marble is black?

Here are two arguments that lead to two different answers for Problem 3. Your job is to figure out which is wrong.

Argument 3.1. Before knowing the removed marble is black, we have 4 equally likely cases:

1. The original marble was white and it was removed from the bag.
2. The original marble was white and the added black marble was removed.
3. The original marble was black and it was removed from the bag.
4. The original marble was black, but the added black marble was the one removed.

Since we know the removed marble is black, we can eliminate case 1 entirely, leaving 2 out of the 3 remaining cases with a black marble initially, giving us a probability of $\frac{2}{3}$.

Argument 3.2. It's obvious that if the original marble is white, then the remaining marble will be white (since it wasn't removed), and if the original was black, the remaining must be black (both are black). Thus, the colour of the marble in the bag is not changed by adding or removing a black marble. Therefore the probability hasn't changed from the start, and should in fact be $\frac{1}{2}$.

So here we have 2 contradictory results! One must be wrong, surely? Think about it carefully before reading the answer.

Still too easy? Consider two other variants:

Problem 4. You looked into the bag from Problem 3 and removed a black marble deliberately.

Problem 5. You picked a marble from the same bag from Problem 3, but you are blind so you don't know what the colour of the removed marble is.

So what's the probability that this article is good? Probably zero. But we leave that as an exercise to the reader.

Solutions to the problems

1. $\frac{2}{7}$. Consider if we chose them in the order RBG. We get the probability to be $\frac{2}{9} \times \frac{3}{8} \times \frac{4}{7} = \frac{1}{21}$. Notice that if we chose them in a different order, we would still obtain the same probability. Hence, our total probability is just the sum of all 6 orders in which we can choose the balls, hence its $\frac{1}{21} \times 6 = \frac{2}{7}$.

2. $\frac{5}{7}$. Leap years can have 52 or 53 Sundays, with 366 days out of which there are 52 complete weeks and 2 remaining (consecutive) days. There are 7 cases for these remaining days, out of which 2 yield 53 Sundays (being (Sat, Sun) and (Sun, Mon)); hence the remaining 5 will yield 52 Sundays. Hence the probability is $\frac{5}{7}$.

3. $\frac{2}{3}$. At first, Argument 3.2 might seem plausible, but it is actually Argument 3.1 which is correct, with the probability being $\frac{2}{3}$. The fallacy in Argument 3.2 is in the last sentence, in which it claims the probability hasn't changed by adding and removing a black marble. This is not true, as the probability has changed, since the black marble was chosen at random. By seeing that the removed marble was black, this reduces our possible cases from 4 to 3, and this changes our probability accordingly.

4. $\frac{1}{2}$. This time, Argument 3.2 applies as there is no effect on probability if we add a black marble and then remove it.

5. $\frac{3}{4}$. Without information about the removed marble, any of the 4 cases in Argument 3.1 are possible. With 3 of the 4 cases having a black marble remaining, our probability is $\frac{3}{4}$.

As I have no better place to put this, here's another unrelated problem: What's the probability at least two NUSH students have the same number of friends in NUSH? This is also left as an exercise to the reader.

References

Smullyan, R. (2015). *The Magic Garden of George B and Other Logic Puzzles.* World Scientific, pp. 5–7.

Existence and uniqueness results for the reduced row echelon form

Ho Boon Suan

Row reduction is a nice systematic way of solving systems linear equations; things like $3x + 4y = 5$. It's a key tool in linear algebra, which is a powerful branch of mathematics that shows up everywhere you look in math, science and engineering. For a beautiful overview, I refer you to 3Blue1Brown's *Essence of Linear Algebra* series on YouTube.

I couldn't really find any treatment of this topic that I liked, so I made my own. Keep in mind that *none of this work is original*—it's just for me to collect the various related proofs in one place. Also, this article is aimed at people who already know the basics of linear algebra. We begin with a few definitions.

Definition 1. Let A be an $m \times n$ matrix. Any one of the following three operations on the rows of A is called an **elementary row operation**[22]:

1. interchanging any two rows of A;
2. multiplying any row of A by a nonzero scalar;
3. adding any scalar multiple of a row of A to another row.

[22] The choice of row operations instead of column operations is arbitrary—one could have an analogous discussion for column operations.

Elementary row operations are of **type 1**, **type 2**, or **type 3** depending on whether they are obtained by (1), (2), or (3). If one may obtain the result of an elementary row operation via the left-multiplication of a matrix, that matrix is said to be **elementary**. Such matrices are invertible.

For example, multiplication on the left by the elementary matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

corresponds to a type 1 operation that switches rows 1 and 2.

Definition 2. A matrix is said to be in **reduced row echelon form** if the following three conditions are satisfied.

a. Any row containing a nonzero entry precedes any row in which all the entries are zero (if any).

b. The first nonzero entry in each row is the only nonzero entry in its column. We call such entries **pivots**.

c. The first nonzero entry in each row is 1 and it occurs in a column to the right of the first nonzero entry in the preceding row.

For example, the matrix

$$\begin{pmatrix} 1 & 2 & 0 & 4 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

is in reduced row echelon form, but the matrix

$$\begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

is not, as it fails to satisfy condition (b).

Theorem 1. *Any matrix can be transformed into reduced row echelon form by finitely many elementary row operations.*

Proof.[23] We proceed by induction on rows. Let A be a matrix with one row. If A is the zero matrix, then it is already in reduced row echelon form, so we are done. If A is non-zero, then there exists some non-zero entry A_{1j} such that all the entries to its left are zero. Multiplying the first and only row of A by $1/A_{1j}$, the matrix ends up in reduced row echelon form.

[23] Milan Vujičić. *Linear Algebra Thoroughly Explained*. Springer, 2008

Now, suppose that any matrix with k rows can be transformed into reduced row echelon form by finitely many elementary row operations. Let A be a $(k+1) \times n$ matrix. If A is the zero matrix, then we are done. Suppose A is non-zero. Then there exists some non-zero entry A_{st} such that $A_{ij} = 0$ whenever $j < t$. (That is, the t^{th} column is the first non-zero column.)

$$A = \begin{pmatrix} 0 & \cdots & 0 & * & \cdots & * \\ 0 & \cdots & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & A_{st} & \cdots & * \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & * & \cdots & * \end{pmatrix}$$

If $s \neq 1$, switch rows s and 1 so that the non-zero entry of column t ends up in the first row. Multiplying by $1/A_{st}$, we end up with

$$A \to \begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & * & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & * & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & * & * & \cdots & * \end{pmatrix}.$$

Now, by subtracting appropriate multiples of the first row, we can clear out the rest of the t^{th} column to get

$$A \to \left(\begin{array}{ccccccc} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & * & \cdots & * \end{array} \right).$$

Notice that the outlined submatrix of A at the bottom right has k rows. By the inductive hypothesis, we can bring it into reduced row echelon form via a sequence of elementary row operations. Applying the same operations to the bottom k rows of our matrix, we obtain a matrix

$$A \to \begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & & & \\ \vdots & & \vdots & \vdots & & & \\ 0 & \cdots & 0 & 0 & & & \\ \vdots & & \vdots & \vdots & & & \\ 0 & \cdots & 0 & 0 & & & \end{pmatrix}$$

where the outlined submatrix is in reduced row echelon form. Finally, in all columns of the submatrix where there is a pivot, we subtract appropriate multiples of the row containing the pivot from the first row so that the column only contains one non-zero entry (the pivot).

Since everything to the left to the outlined submatrix is zero, conditions (a), (b) and (c) follow, so that A can be transformed into reduced row echelon form by finitely many elementary row operations. □

We have shown that any matrix has a reduced row echelon form, but it remains to be shown that it is unique.[24]

[24] The *rank* of a matrix measures how much of a degenerate your matrix is; see 3Blue1Brown's videos, in particular https://www.youtube.com/watch?v=uQhTuRlWMxw.
Also, $\{e_1, \ldots, e_n\}$ represents the standard basis for $\mathbf{R}^n$ (just like how $\hat{\imath}$, $\hat{\jmath}$ and $\hat{k}$ form a basis of 3D space $\mathbf{R}^3$).

Theorem 2. *Let A be an $m \times n$ matrix of rank r, where $r > 0$, and let R be a reduced row echelon form of A. Then*

a. The number of nonzero rows in R is r.

b. For each $i = 1, \ldots, r$, there is a column R_{j_i} of R such that $R_{j_i} = e_i$.

c. The columns of A numbered $j_1, \ldots, j_r$ are linearly independent.

d. For each $k = 1, \ldots, n$, if column k of R is $d_1 e_1 + \cdots + d_r e_r$, then column k of A is $d_1 A_{j_1} + \cdots + d_r A_{j_r}$.

Proof.[25] We denote the columns of A and R as $A_1, \ldots, A_n$ and $R_1, \ldots, R_n$ respectively. If the rank of A is r, then R also has rank r as elementary row operations preserve row rank. Because R is in reduced row echelon form, the rows of R are linearly independent.[26] Hence R must have exactly r rows, proving (a).

[25] S. H. Friedberg, A. J. Insel, and L. E. Spence. *Linear Algebra*. Pearson, 2003

[26] This is because each pivot of a row in R is strictly to the right of the pivot of the row above it, and pivot entries are the only non-zero entries in their columns. Thus the only way to get a zero in any column is by taking the trivial linear combination of rows.

If $r \geq 1$, then it follows from the defintion of reduced row echelon form that the vectors $e_1, \ldots, e_r$ must occur among the columns of R. That is, for each $i = 1, \ldots, r$, there is a column R_{j_i} of R such that $R_{j_i} = e_i$, proving (b).

We claim that $A_{j_1}, \ldots, A_{j_r}$ are linearly independent. For suppose there are scalars $c_1, \ldots, c_r$ such that

$$c_1 A_{j_1} + \cdots + c_r A_{j_r} = 0.$$

Because R is obtained from A by left-multiplication by a finite sequence of invertible elementary matrices, we have that $R = MA$ for some invertible $m \times m$ matrix M. Thus we can multiply the above equation by M to get

$$c_1 M A_{j_1} + \cdots + c_r M A_{j_r} = 0.$$

Since $MA_{j_i} = R_{j_i} = e_i$, it follows that

$$c_1 e_1 + \cdots + c_r e_r = 0.$$

Hence $c_1 = \cdots = c_r = 0$, so that the vectors $A_{j_1}, \ldots, A_{j_r}$ are independent, proving (c).

Finally, because R has only r non-zero rows, every column of R has the form $d_1 e_1 + \cdots + d_r e_r$ for scalars $d_1, \ldots, d_r$, so that the corresponding column of A must be

$$\begin{aligned} M^{-1}(d_1 e_1 + \cdots + d_r e_r) &= d_1 M^{-1} e_1 + \cdots + d_r M^{-1} e_r \\ &= d_1 M^{-1} R_{j_1} + \cdots + d_r M^{-1} R_{j_r} \\ &= d_1 A_{j_1} + \cdots + d_r A_{j_r}, \end{aligned}$$

proving (d). □

Corollary. *The reduced row echelon form of a matrix is unique.*

Proof. Let A be an $m \times n$ matrix. If A is the zero matrix, then every elementary row operation on A leaves A unchanged, so that the reduced row echelon form of A is itself, and is thus unique.

Suppose then that A is non-zero with rank $r > 0$. Let R be a reduced row echelon form of A. Then, $R = MA$ for some invertible $m \times m$ matrix M (that is the product of elementary matrices). Let $A_1, \ldots, A_n$ and $R_1, \ldots, R_n$ denote the columns of A and R respectively. By Theorem 2(b), e_i is a column of R for $1 \leq i \leq r$. Let j_i denote the column number of the leftmost column of R such that $R_{j_i} = e_i$. By Theorem 2(c), $\{A_{j_1}, \ldots, A_{j_r}\}$ is linearly independent and therefore is a basis for the column space of A. So, for $1 \leq k \leq n$, we have $A_k = c_1 A_{j_1} + \cdots + c_r A_{j_r}$ for unique scalars $c_1, \ldots, c_r$. Thus

$$\begin{aligned} R_k &= MA_k \\ &= c_1 MA_{j_1} + \cdots + c_r MA_{j_r} \\ &= c_1 R_{j_1} + \cdots + c_r R_{j_r} \\ &= c_1 e_1 + \cdots + c_r e_r \end{aligned}$$

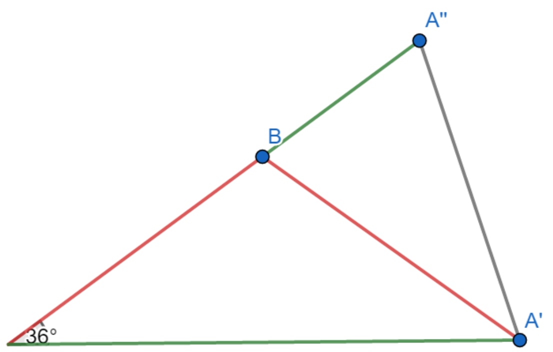
for $1 \leq k \leq n$. Hence the columns of R are completely (and thus uniquely) determined by the columns on A. □

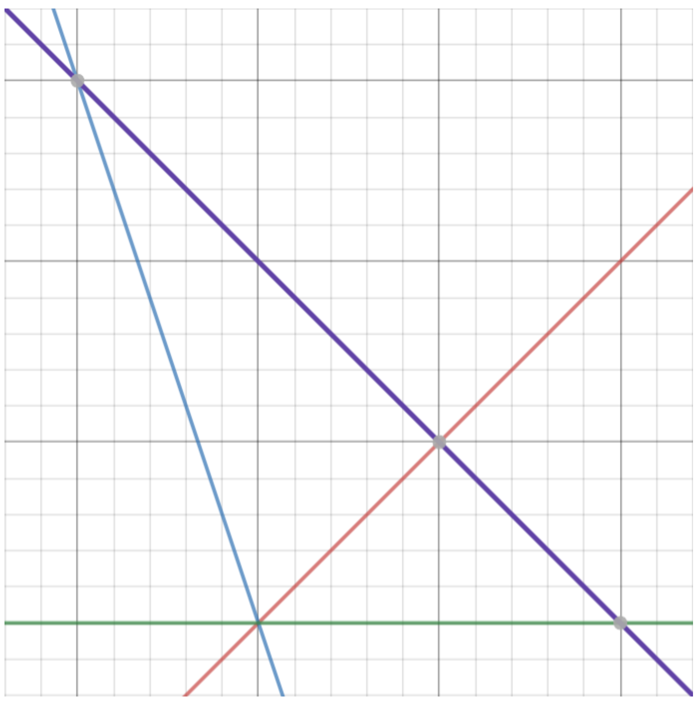
A picture speaks a thousand words

Zhang Xiaorui

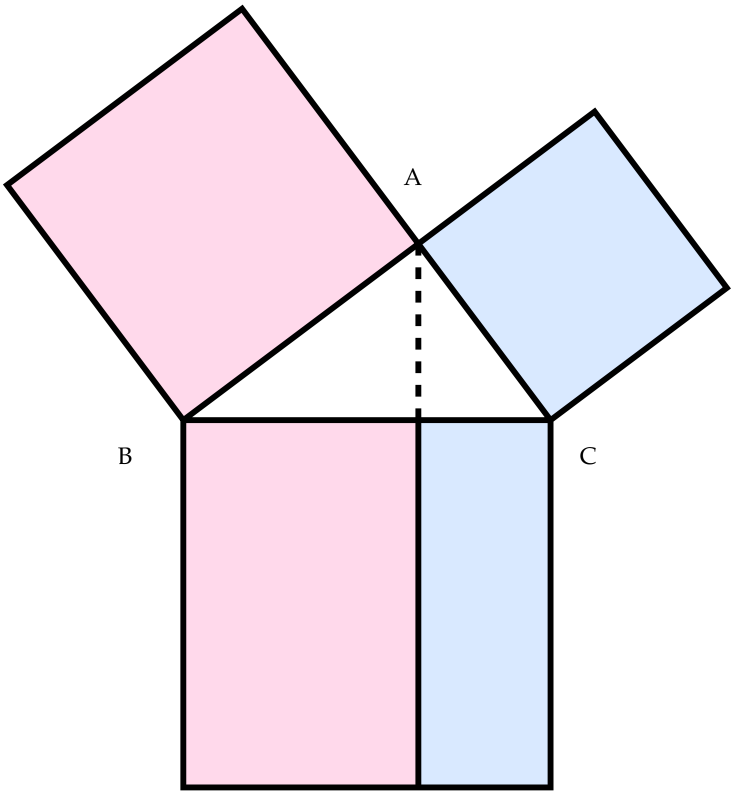
Arranged by difficulty, each of the following pictures tells a story or illustrates some interesting mathematical concepts. This article serves then as an open-ended exploration of these ideas. Play around with these diagrams and pictures and see what you find! After all, mathematics is understood through tinkering and discovery. If you're stumped, there are brief and perhaps cryptic hints at the end. Have fun!

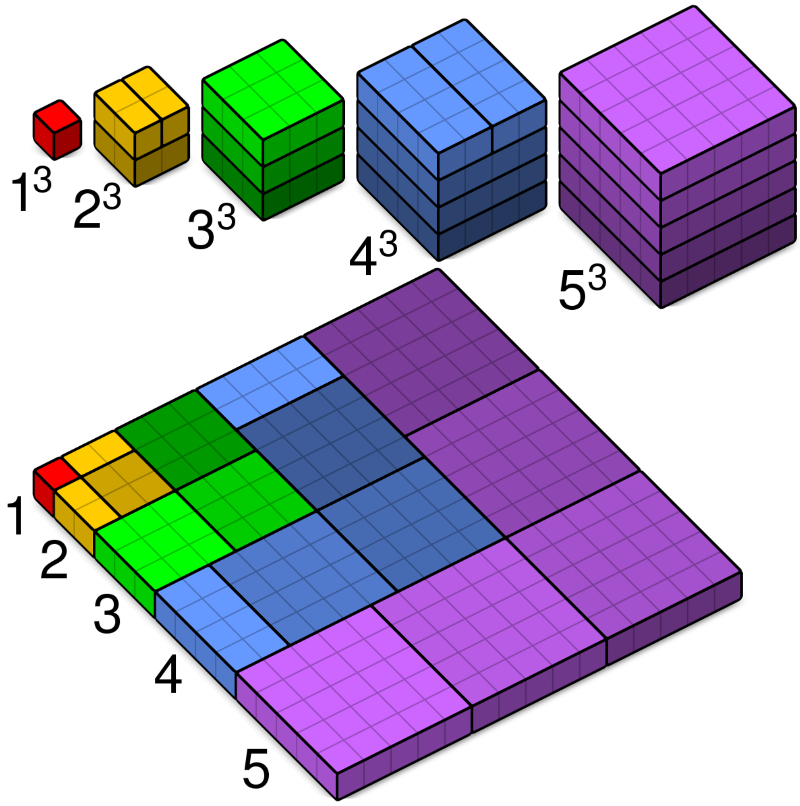
Simple



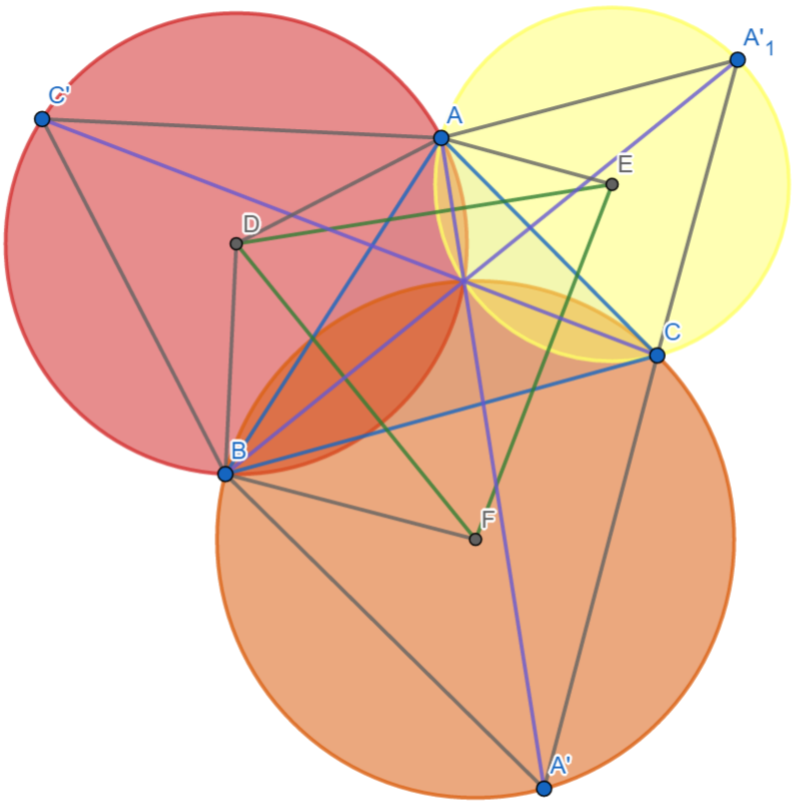


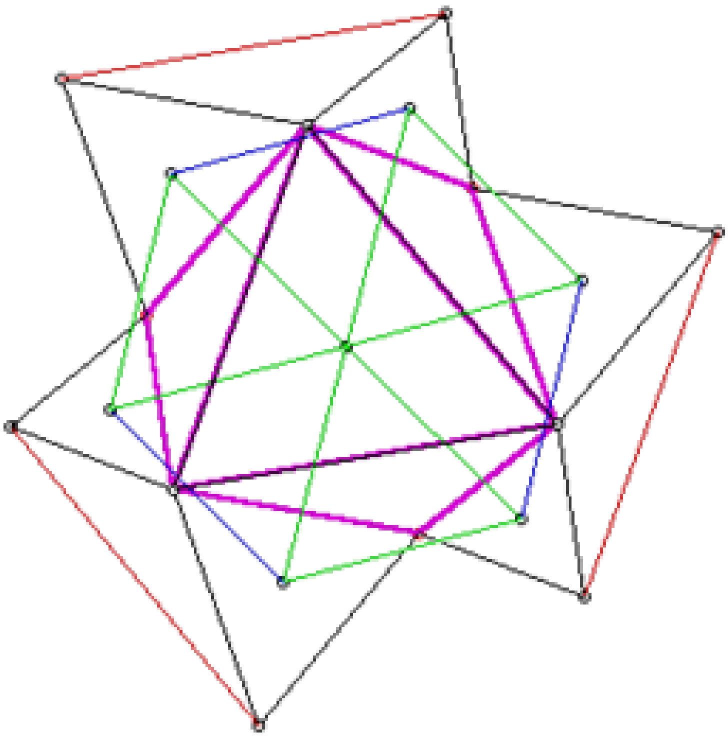
Medium



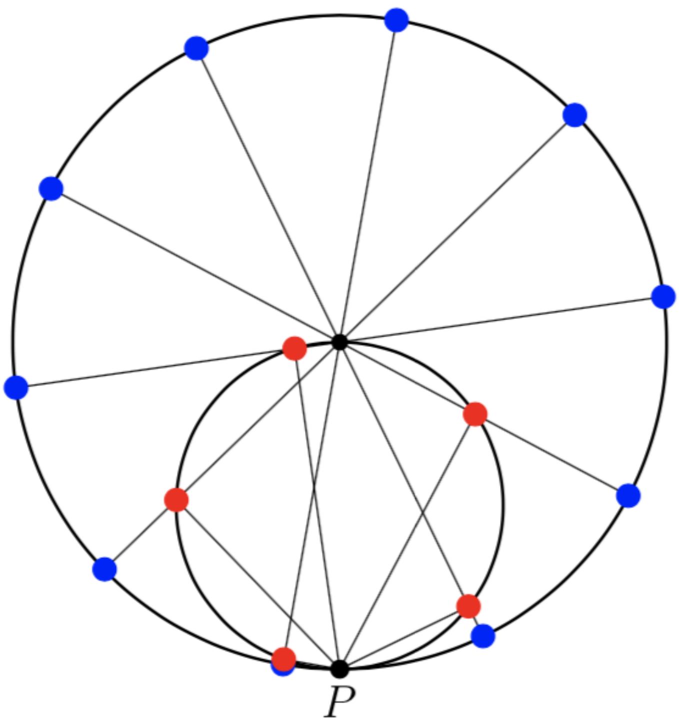


Challenging





Bonus



$$\frac{\text{🍎}}{\text{🍌} + \text{🍐}} + \frac{\text{🍐}}{\text{🍌} + \text{🍎}} + \frac{\text{🍌}}{\text{🍎} + \text{🍐}} = 4$$

$$(\text{🍐}, \text{🍎}, \text{🍌}) \in \mathbb{N}^3, \text{🍌} =?, \text{🍎} =?, \text{🍐} =?$$

Hints are provided on the next page.

Hints

Simple

Top: Fibonacci
Bottom: Tangents

Medium

Top: Pythagoras, but by what?
Bottom: Sums

Challenging

Top: This is a bit confusing, it's proving 3 different things at once. Try to identify them.
Bottom: Proving the same theorem, but using a rather uncommon method.

Bonus

Top: _B_B: "Where's the circle?"
Bottom: 99% of NUS High Students can't solve this! Can you find positive whole values for the apple, pear and banana?

Credits

I HAVE DONE MY BEST to properly attribute any intellectual property used within this issue of the MIG magazine below. If I have made any mistakes, I sincerely apologize.

— THE EDITOR

Page 1. *Squaring the circle*. Wikimedia Commons, Public Domain. Original PNG by `Plynn9`; SVG by Alexei Kouprianov. https://commons.wikimedia.org/wiki/File:Squaring_the_circle.svg
Page 1. *Be rational; get real*. Designed by `alexlin9008` on Zazzle. https://www.zazzle.com/be_rational_get_real_poster-228943564713108078
Page 4. *Square wave*. Adapted from Gonzalo Medina's answer at TEX StackExchange. https://tex.stackexchange.com/a/113051
Page 18. *Illustration to Euclid's proof of the Pythagorean theorem*. Wikimedia Commons, WTFPL version 2. Original SVG by `Gerbrant`. https://commons.wikimedia.org/wiki/File:Illustration_to_Euclid's_proof_of_the_Pythagorean_theorem.svg
Page 18. *Visual proof of Nicomachus' theorem*. Wikimedia Commons, Creative Commons Attribution-Share Alike 3.0 Unported.
Original SVG by `Cmglee`. https://commons.wikimedia.org/wiki/File:Nicomachus_theorem_3D.svg
Page 19. *Napoleon's Theorem, Two Simple Proofs*. Scott Brodie. https://www.cut-the-knot.org/proofs/napoleon.shtml
Page 20. Figure 3 in *Circles, the Basel Problem, and the Apparent Brightness of Stars*. Matt Baker's Math Blog. https://mattbaker.blog/2018/03/14/circles-the-basel-problem-and-the-apparent-brightness-of-stars/

THE END