Experimental Mathematics 2020

Alex Ghitza

Version of Tue 31st Mar, 2020 at 14:13

1 Gauss's agM **2**
1.1 Two sequences, one limit . . . 2
1.2 An elliptic integral . . . 4
1.3 Built-in agM commands . . . 7

2 Introduction to Sage and Mathematica **8**
2.1 Lines on a plane . . . 8
2.2 Pascal mod 2 . . . 10

3 Constant recognition **13**
3.1 Finding the fraction . . . 13
3.2 Finding the integer relation . . . 16
3.3 Lattice reduction . . . 17

Some answers/solutions/hints/more questions **20**

1 Gauss's agM

From Gauss's diary, entry dated 30 May 1799:

> Terminum medium arithmetico–geometricum inter 1 et $\sqrt{2}$ esse = $\frac{\pi}{\varpi}$ usque ad figuram undecimam comprobavimus, qua re demonstrata prorsus novus campus in analysis certo aperietur.

For those as Latin-challenged as I am, this translates to

> We have established that the arithmetic–geometric mean between 1 and $\sqrt{2}$ is = $\frac{\pi}{\varpi}$ to the eleventh decimal place; the demonstration of this fact will surely open an entirely new field of analysis.

This probably raises more questions than it answers. So let's look at it in more detail.

1.1 Two sequences, one limit

Consider the two sequences of real numbers (a_n) and (b_n) defined by the recursions

$$a_0 = \sqrt{2}, \qquad a_{n+1} = \frac{a_n + b_n}{2}$$
$$b_0 = 1, \qquad b_{n+1} = \sqrt{a_n b_n}$$

This looks a little strange, as the two definitions are intertwined (and what's up with the initial values $\sqrt{2}$ and 1?), but we can recognise some familiar features, at least locally. The recursive rule defining a_{n+1} takes the arithmetic mean of the numbers a_n and b_n. And the recursive rule defining b_{n+1} takes the geometric mean of the same two numbers.

I wonder how the two sequences behave as n increases. If I were Gauss, I would compute the first few terms by hand (to 11 decimal places, indeed). Having the luxury of living in the era of funny cat YouTube videos (and incidental infrastructure), I'll instead call upon SageMath on my computer:

```
sage: def a(n):
....:     if n == 0:
....:         return RR(sqrt(2))
....:     else:
....:         return RR((a(n-1)+b(n-1))/2)
sage: def b(n):
....:     if n == 0:
....:         return RR(1)
....:     else:
....:         return RR(sqrt(a(n-1)*b(n-1)))
```

Having defined the two sequences, I can now ask for values:

```
sage: a(1)
1.20710678118655
sage: a(2)
1.19815694809463
sage: a(3)
1.19814023479388
sage: a(4)
1.19814023473559
sage: a(5)
1.19814023473559
```

One might guess now that (a_n) converges to a number close to 1.19814023473559. How about (b_n)?

```
sage: b(1)
1.18920711500272
sage: b(2)
1.19812352149312
sage: b(3)
1.19814023467731
sage: b(4)
1.19814023473559
sage: b(5)
1.19814023473559
```

This also seems to converge, and to the same limit? Let's be bold and proclaim:

Proposition 1.1. *Given any starting values $a_0 = a \geq b_0 = b \in \mathbb{R}_{>0}$, the sequences (a_n) and (b_n) both converge to the same limit $M(a, b)$.*

Proof. More of a sketch, really.

There are two things to prove: that the sequences converge, and that they have the same limit[1].

Here is one approach to this, courtesy of Wikipedia:

- $b_n \leq a_n$ for all n. This is clear for $n = 0$ and true by simple algebraic manipulation in general (the arithmetic mean of two numbers is never smaller than their geometric mean).
- Use the previous part to note that $b_{n+1} \geq b_n$, so the sequence (b_n) is non-decreasing.
- Note that $b_n \leq a$ for all $n \geq 0$, so the sequence (b_n) is bounded above. Hence it has a limit L.
- Note that $b_n \geq b > 0$ for all n so $L \geq b > 0$, in particular $L \neq 0$.
- Note that
$$a_n = \frac{b_{n+1}^2}{b_n},$$
so using a theorem about the limit of the quotient of two convergent sequences we conclude that (a_n) converges to $\frac{L^2}{L} = L$.

[1]It would not be sufficient to only show that the sequence of differences $(a_n - b_n)$ converges to 0

□

The real number $M(a, b)$ is called the *arithmetic-geometric mean (agM) of a and b*. The above numerical experiment leads us to believe that

$$M(\sqrt{2}, 1) = 1.19814023473559\ldots$$

But Gauss's diary entry said more about this value. What's that about?

1.2 An elliptic integral

Consider the *lemniscate (of Bernoulli)*, a plane curve given in polar coordinates (r, θ) by the equation

$$r^2 = \cos(2\theta).$$

It would be nice to graph it, wouldn't it? As far as I know, neither Sage nor Mathematica have a built-in command for implicit polar plots. After a bit of algebraic manipulation with $x = r\cos\theta$, $y = r\sin\theta$, I get the implicit Cartesian equation

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

This is more amenable to plotting:

```
sage: var('x, y')
(x, y)
sage: f = (x^2 + y^2)^2 - (x^2 - y^2)
sage: p = implicit_plot(f, (x, -1, 1), (y, -1, 1))
sage: p.show()
None
```

which produces the pretty infinity sign in the picture.

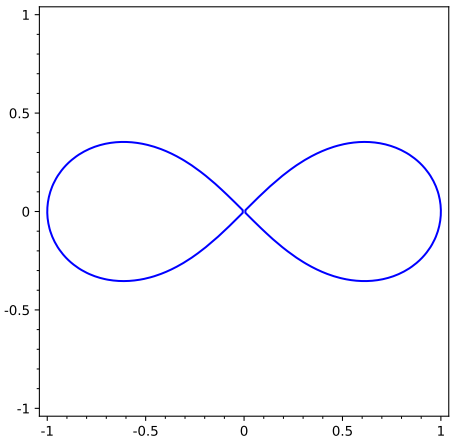


Figure 1.1: The lemniscate of Bernoulli, from Sage

Or you can follow a Mathematica lead from StackExchange:

https://mathematica.stackexchange.com/a/549

```
ContourPlot[
 Evaluate@With[
   {r = Sqrt[x^2 + y^2],
    theta = ArcTan[x, y]},
   r^2 - Cos[2*theta] == 0
   ],
{x, -1, 1}, {y, -1, 1}
]
```

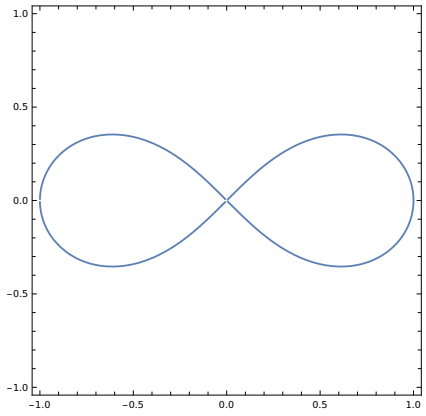


Figure 1.2: The lemniscate of Bernoulli, Mathematica style

The question is: what is the arclength of this curve?
The standard arclength formula in polar coordinates gives

$$4\int_0^{\pi/4}\left(r^2+\left(\frac{dr}{d\theta}\right)^2\right)^{1/2}d\theta = 4\int_0^{\pi/4}(\cos(2\theta))^{-1/2}\,d\theta.$$

If we introduce a new variable α with the property that $\cos(2\theta) = \cos^2\alpha$, the integral becomes

$$4\int_0^{\pi/2}\left(1+\cos^2\alpha\right)^{-1/2}d\alpha = 4\int_0^{\pi/2}\left(2\cos^2\alpha+\sin^2\alpha\right)^{-1/2}d\alpha.$$

One-half of this value is what Gauss denoted by ϖ in his diary entry.
We shouldn't just believe him like this. Let's check his calculations by estimating the integral numerically.
In Sage:

```
sage: f = 2/sqrt(2*cos(x)^2 + sin(x)^2)                                    37
sage: varpi = numerical_integral(f, 0, pi/2)[0]                            38
sage: RR(pi/varpi)                                                         39
1.19814023473559                                                           40
```

Or in Mathematica:

```
In[1]:= Pi/NIntegrate[2/Sqrt[2*Cos[t]^2+Sin[t]^2],{t,0,Pi/2},
                      {WorkingPrecision->15}]

Out[1]= 1.19814023473559
```

Fine, Gauss seems to have gotten his decimals right. More generally, he proved

Theorem 1.2. *If $a \geq b > 0$ then*

$$\int_0^{\pi/2} \left(a^2 \cos^2 \alpha + b^2 \sin^2 \alpha\right)^{-1/2} \, d\alpha = \frac{\pi}{2M(a,b)}$$

(These are examples of so-called *elliptic integrals*, and the arithmetic-geometric mean is a very efficient way of approximating them.)

Proof. Write

$$I(a,b) = \int_0^{\pi/2} \left(a^2 \cos^2 \alpha + b^2 \sin^2 \alpha\right)^{-1/2} \, d\alpha$$

We introduce a variable φ with the property that

$$\sin \alpha = \frac{2a \sin \varphi}{a + b + (a-b) \sin^2 \varphi} \tag{1.1}$$

Then some secret magic sauce (in Gauss's words *"after the development has been done correctly, it will be seen"*[2]):

$$\left(a^2 \cos^2 \alpha + b^2 \sin^2 \alpha\right)^{-1/2} \, d\alpha = \left(a_1^2 \cos^2 \varphi + b^2 \sin^2 \varphi\right)^{-1/2} \, d\varphi$$

which leads us to the conclusion

$$I(a,b) = I(a_1, b_1).$$

A moment's further thought brings us to continue this as

$$I(a,b) = I(a_1,b_1) = I(a_2,b_2) = \cdots = I(a_n,b_n) = \ldots$$

and some judicious (real) analysis allows us to pass to the limit and get

$$I(a,b) = I(M(a,b), M(a,b)) = \frac{\pi}{2M(a,b)}$$

as the integral $I(c,c)$ is a piece of cake. □

Want to know more? There is plenty more where this came from, namely David Cox's articles [Cox85] (the shorter version) and [Cox84] (the longer one).

Ah, I can't resist mentioning one more thing, which I learned from Cox's papers. In 1973, Salamin discovered the formula

$$\pi = \frac{2M(\sqrt{2},1)^2}{1 - \sum_{n=1}^{\infty} 2^{n+1}(a_n^2 - b_n^2)}$$

which, as you can see, uses an infinite sum involving the terms a_n and b_n of the sequence we started with, as well as the common limit $M(\sqrt{2},1)$ of these two sequences. The existence of such a formula for π is weird enough, but it turns out [Sal76] that it's actually a pretty efficient way to compute lots of digits of π (in that the number of significant digits doubles at each step).

[2] Jacobi must have been as unimpressed as us by Gauss weaseling out of this, because he wrote down a couple of intermediate steps: Given the relation (1.1) between α and φ, show that

$$\cos \alpha = \frac{(2 \cos \varphi)\left(a_1^2 \cos^2 \varphi + b_1^2 \sin^2 \varphi\right)^{1/2}}{a + b + (a-b) \sin^2 \varphi}$$

and then that

$$\left(a^2 \cos^2 \alpha + b^2 \sin^2 \alpha\right)^{1/2} = a \, \frac{a + b - (a-b) \sin^2 \varphi}{a + b + (a-b) \sin^2 \varphi}$$

After this and implicitly differentiating (1.1), it's all smooth sailing to Gauss's equality of differentials.

1.3 Built-in agM commands

You might wonder if the arithmetic-geometric mean is already implemented in the software we're playing with.

But of course! In Mathematica, it is as simple as `ArithmeticGeometricMean`.

In Sage, it is a method (i.e. function attached to an object) of elements of `RR` (or other `RealField`s, and variants):

```
sage: a = RR(sqrt(2))
sage: a.agm(1)
1.19814023473559
```

Exercise 1.3. Implement your own function `myagm(a,b)` that returns the agM of two positive real numbers a and b. If you are so inclined, think about numerical issues.

2 Introduction to Sage and Mathematica

We look at a couple of innocent questions as an excuse to familiarise ourselves with the basics of the software. I would recommend reading through this in parallel with Sage's guided tour:

https://doc.sagemath.org/html/en/tutorial/tour.html

and bits of Mathematica's fast introduction for math students:

https://www.wolfram.com/language/fast-introduction-for-math-students/en/

2.1 Lines on a plane

What is the maximal number of regions $R(n)$ that you can obtain by drawing n lines in the plane $\mathbb{R}^2$?

We'll explore this by hand first. Clearly

$$R(0) = 1, \qquad R(1) = 2, \qquad R(2) = 4.$$

This is a good point to try guessing (a) the next number $R(3)$ in the sequence and (b) a formula for the general term of the sequence. Popular choices are $R(3) = 8$ and $R(n) = 2(n+1)$ or $R(n) = 2^n$. This is actually incorrect, as paper-and-pen will convince you that

$$R(3) = 7$$

As you continue exploring this by hand, you may notice that, given an existing configuration of lines, a new line creates k new regions if and only if it crosses $k-1$ old lines. So if we want to maximise k then we need to maximise the number of old lines we cross. This leads to the recursion

$$\begin{aligned} R(0) &= 1 \\ R(n) &= R(n-1) + n \qquad \text{if } n > 1 \end{aligned}$$

We can implement this in Sage as

```
sage: def R(n):
....:     if n == 0:
....:         return 1
....:     return R(n-1) + n
```

Let's check the first few terms:

```
sage: R(1)
2
sage: [R(n) for n in range(4)]
[1, 2, 4, 7]
```

This matches what we already knew. And now, into the great unknown:

```
sage: [R(n) for n in range(20)]
[1, 2, 4, 7, 11, 16, 22, 29, 37, 46, 56, 67, 79, 92, 106, 121, 137, 154, 172, 191]
```

In Mathematica, we could do something like

```
In[1]:= Clear[R]
In[2]:= R[0] = 1;
In[3]:= R[n_] := R[n-1] + n
In[4]:= Table[R[n], {n, 0, 19}]

Out[4]= {1, 2, 4, 7, 11, 16, 22, 29, 37, 46, 56, 67, 79, 92, 106, 121, 137,
>    154, 172, 191}
```

Very satisfying. How about a formula for $R(n)$ though? We could think about it and figure it out fairly quickly in this example, but let's instead search for the first few terms in the Online Encyclopedia of Integer Sequences (OEIS):

oeis.org

Feeding it $1, 2, 4, 7, 11, 16, 22, 29$, it gives a number of suggestions, the foremost of which is

> A000124 Central polygonal numbers (the Lazy Caterer's sequence): n(n+1)/2 + 1; or, maximal number of pieces formed when slicing a pancake with n cuts.

In fact, this is such a common workflow (work out the first few terms, consult OEIS), that you can do it entirely from Sage[1]:

```
sage: lst = [R(n) for n in range(8)]
sage: lst
[1, 2, 4, 7, 11, 16, 22, 29]
sage: oeis(lst)
0: A000124: Central polygonal numbers (the Lazy Caterer's sequence): n(n+1)/2 + 1;
    or, maximal number of pieces formed when slicing a pancake with n cuts.
1: A152947: a(n) = 1 + (n-2)*(n-1)/2.
2: A098574: a(n) = Sum_{k=0..floor(n/7)} C(n-5*k,2*k).
```

OEIS helpfully gives us a closed form for the number of regions:

$$R(n) = \frac{n(n+1)}{2} + 1$$

Of course!

$$\begin{aligned}
R(0) &= 1 \\
R(1) &= R(0) + 1 = 1 + 1 \\
R(2) &= R(1) + 2 = 1 + 1 + 2 \\
R(3) &= R(2) + 3 = 1 + 1 + 2 + 3 \\
R(n) &= R(n-1) + n = 1 + \sum_{k=1}^{n} k = 1 + \frac{n(n+1)}{2}
\end{aligned}$$

By the way, in case you forgot the sum of the first n positive integers, both Sage and Mathematica can help:

[1] Check out `oeis?` for more information.

```
sage: k, n = var("k, n")
sage: sum(k, k, 1, n)
1/2*n^2 + 1/2*n
sage: sum(k, k, 1, n).factor()
1/2*(n + 1)*n
```

```
In[1]:= Sum[k, {k, 1, n}]

        n (1 + n)
Out[1]= ---------
            2
```

Exercise 2.1. Go one dimension up and play with it: what is the maximal number of regions that can be obtained from n planes in $\mathbb{R}^3$?

As frivolous as the topic may seem (slicing a pancake with n cuts, indeed), it is an active area of research. Look up hyperplane arrangements on the web, or [Sta12, Section 3.11].

2.2 Pascal mod 2

What is the distribution of even and odd numbers in Pascal's triangle?

Let's start by generating part of Pascal's triangle. The glorious way to approach this is to code the recursive construction of the triangle, but we're after answers rather than glory so we'll just use the built-in `binomial` to get the binomial coefficients:

```
sage: lst = [binomial(m, n) for m in range(2^3) for n in range(2^3)]
sage: lst
[1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 1, 2, 1, 0, 0, 0, 0, 0, 1, 3, 3,
   1, 0, 0, 0, 0, 1, 4, 6, 4, 1, 0, 0, 0, 1, 5, 10, 10, 5, 1, 0, 0, 1, 6, 15, 20,
   15, 6, 1, 0, 1, 7, 21, 35, 35, 21, 7, 1]
sage: matrix(2^3, lst)
[ 1  0  0  0  0  0  0  0]
[ 1  1  0  0  0  0  0  0]
[ 1  2  1  0  0  0  0  0]
[ 1  3  3  1  0  0  0  0]
[ 1  4  6  4  1  0  0  0]
[ 1  5 10 10  5  1  0  0]
[ 1  6 15 20 15  6  1  0]
[ 1  7 21 35 35 21  7  1]
```

Looks promising. Let's reduce modulo 2:

```
sage: lst = [binomial(m, n) % 2 for m in range(2^3) for n in range(2^3)]
sage: matrix(2^3, lst)
[1 0 0 0 0 0 0 0]
[1 1 0 0 0 0 0 0]
[1 0 1 0 0 0 0 0]
[1 1 1 1 0 0 0 0]
[1 0 0 0 1 0 0 0]
[1 1 0 0 1 1 0 0]
[1 0 1 0 1 0 1 0]
[1 1 1 1 1 1 1 1]
```

We can maybe see a pattern, but a bigger version might help. And dots are more easily visualised than 0s and 1s. So let's plot a dot whenever the binomial coefficient is odd:

```
sage: lst = [(m, n) for m in range(2^8) for n in range(2^8) \
....:        if (binomial(m, n) % 2) == 1]
sage: p = list_plot(lst)
sage: p.show()
None
```

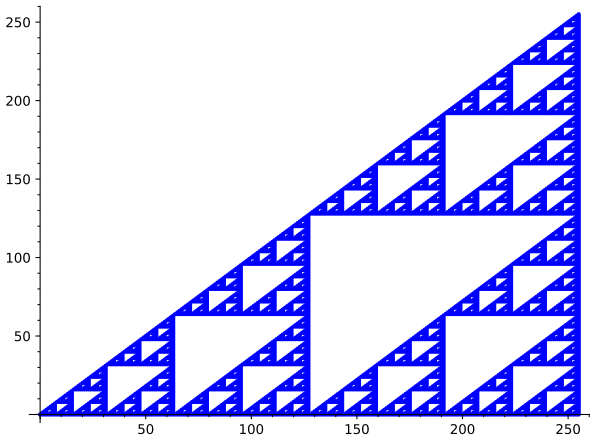


Figure 2.1: Pascal's triangle modulo 2

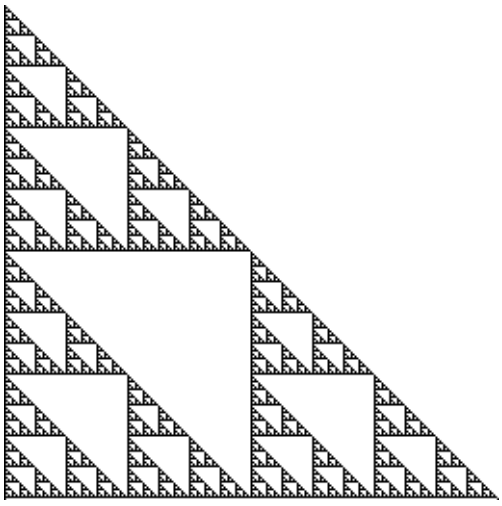
For a proof of what you are observing visually (i.e. that the picture looks a lot like Sierpiński's gasket), see [Ste95, Encounter 2].

The Mathematica version:

```
In[1]:= Table[Binomial[n, m], {n, 0, 2^3}, {m, 0, 2^3}] // MatrixForm

Out[1]//MatrixForm= 1    0    0    0    0    0    0    0    0

                    1    1    0    0    0    0    0    0    0

                    1    2    1    0    0    0    0    0    0

                    1    3    3    1    0    0    0    0    0

                    1    4    6    4    1    0    0    0    0

                    1    5    10   10   5    1    0    0    0

                    1    6    15   20   15   6    1    0    0

                    1    7    21   35   35   21   7    1    0

                    1    8    28   56   70   56   28   8    1
```

```
In[2]:= Table[Mod[Binomial[n, m], 2], {n, 0, 2^3}, {m, 0, 2^3}] // MatrixForm

Out[2]//MatrixForm= 1   0   0   0   0   0   0   0   0

                    1   1   0   0   0   0   0   0   0

                    1   0   1   0   0   0   0   0   0

                    1   1   1   1   0   0   0   0   0

                    1   0   0   0   1   0   0   0   0

                    1   1   0   0   1   1   0   0   0

                    1   0   1   0   1   0   1   0   0

                    1   1   1   1   1   1   1   1   0

                    1   0   0   0   0   0   0   0   1
```

```
In[3]:= Table[Mod[Binomial[n, m], 2], {n, 0, 2^8}, {m, 0, 2^8}] // Image // ColorNegate
```



Exercise 2.2. Look up how to plot a matrix in Sage. Use this to produce a visualisation of Pascal's triangle mod 2.

Exercise 2.3. Visualise the reduction of Pascal's triangle modulo other integers. Start with 3, 4, 5, ... Maybe give different colours to the different remainders.

3 Constant recognition

A common occurrence in computer-assisted mathematics is obtaining a numerical approximation to a number we are interested in. How can we recognise whether this number is of a special type (e.g. rational, or algebraic, or a simple combination of other numbers we like such as π, e, $M(\sqrt{2},1)$)? There are surprisingly robust ways of approaching such an ill-defined question, as we'll see now.

3.1 Finding the fraction

The Riemann zeta-function is a function of a complex variable s defined, for $\mathrm{Re}(s) > 1$, by the infinite series

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

Despite having Riemann's (1826–1866) name attached to it, particular aspects had been considered two hundred years before Riemann. As one such example, Mengoli asked in 1650 for the value

$$\zeta(2) = \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} = \dots$$

What Mengoli meant by "value" was not an approximation such as

```
sage: RR(zeta(2))                                                    93
1.64493406684823                                                     94
```

which he could compute himself (maybe to slightly fewer decimals). It was the exact value, in closed form, a notion that is more metaphysical than mathematical. He was basically saying "I want a simple and pretty answer involving terms that I know already, and a proof that the answer is correct." This became known as the Basel problem and Euler knocked it out of the ballpark in 1741.

We're going to leave the historical trail and imagine for a moment that, by some stroke of genius, we thought it would be a good idea to divide $\zeta(2)$ by π^2 (why not?)

```
sage: RR(zeta(2)/pi^2)                                               95
0.166666666666667                                                    96
```

That is very compelling. Is it an artifact of the low precision?

```
sage: Rbig = RealField(1000)                                         97
sage: Rbig(zeta(2)/pi^2) == Rbig(1/6)                                98
True                                                                 99
```

Indeed, as Euler proved, it is the case that

$$\frac{\zeta(2)}{\pi^2} = \frac{1}{6}$$

Let's try another one. What could this be:

```
sage: RR(zeta(4))                                                    100
```

```
1.08232323371114
```

Maybe more divine revelation can help:

```
sage: RR(zeta(4)/pi^4)
0.0111111111111111
```

Yes,

$$\frac{\zeta(4)}{\pi^4} = \frac{1}{90}$$

This is easy! I can't believe there's a whole chapter devoted to this. Okay, one more:

```
sage: RR(zeta(12)/pi^12)
1.08220214040320e-6
```

Huh? Surely we need more decimals:

```
sage: Rbig = RealField(250)
sage: a = Rbig(zeta(12)/pi^12)
sage: a
1.0822021404031986042568053150063732074314084896095478106060116642127224138e-6
```

There's a certain gadget called a *continued fraction*, which is tailor-made for our problem[1]:

```
sage: c = continued_fraction(a)
sage: c
[0; 924041, 1, 3, 1, 2, 2, 1, 14]
```

What this really means is

$$\frac{\zeta(12)}{\pi^{12}} = 0 + \cfrac{1}{924041 + \cfrac{1}{1 + \cfrac{1}{3 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{1 + \frac{1}{14}}}}}}}} = \frac{691}{638512875}$$

More generally, a continued fraction expansion of a positive real number β is an expression of the form

$$\beta = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \frac{1}{\cdots}}}}$$

where $a_0 \in \mathbb{Z}_{\geq 0}, a_1, a_2, \cdots \in \mathbb{Z}_{>0}$.

Given β, the continued fraction expansion is easily computed by the recursion

$$\begin{aligned} \beta_0 &= \beta \\ a_n &= \lfloor \beta_n \rfloor \qquad \text{for } n \geq 0 \\ \beta_n &= \frac{1}{\beta_{n-1} - a_{n-1}} \qquad \text{for } n \geq 1 \end{aligned}$$

[1] And it's been around forever. According to Wikipedia, the first documented use of continued fractions is in Sanskrit and goes back to 499.

The truncation at a_n of a continued fraction is called its *n-th convergent*

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}} = \frac{p_n}{q_n}$$

Exercise 3.1. If you harbour any nostalgia for the good ol' Real Analysis days, try your hand at this: prove that the sequence (p_n/q_n) converges as $n \to \infty$. One possible approach is to first show that

$$p_{n-1}q_n - p_n q_{n-1} = (-1)^n$$

then conclude that the sequence (p_n/q_n) is Cauchy.

For our constant-recognition purposes, some relevant facts are

- β is rational if and only if its continued fraction expansion is finite. This is closely related to the Euclidean algorithm for computing the greatest common denominator of two integers.
- β is quadratic irrational if and only if its continued fraction expansion is (eventually) periodic. (A quadratic irrational is a number of the form $a + b\sqrt{d}$ with $a, b \in \mathbb{Q}$ and d a squarefree positive integer.)
- Given a denominator upper bound Q, the convergent p_n/q_n with the largest n such that $q_n \le Q$ is the *best rational approximation to β with denominator at most Q*, in other words it is the fraction p/q that minimises the quantity

$$|q\beta - p| \qquad \text{subject to } 1 \le q \le Q.$$

Some well-known constants have regular-looking continued fractions:

$$\frac{1+\sqrt{5}}{2} = \phi = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{\cdots}}}}}}}}}$$

$$e = 2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{4 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{6 + \cfrac{1}{1 + \cfrac{1}{\cdots}}}}}}}}}$$

while others do not:

$$\pi = 3 + \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{1 + \cfrac{1}{292 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{\ddots}}}}}}}}}}$$

3.2 Finding the integer relation

Let's take some of the discussion in the previous section and twist it around a little bit. Suppose I consider the real numbers $\zeta(2)$ and π^2 and I ask: are there integers a and b such that

$$a\,\zeta(2) + b\,\pi^2 = 0? \tag{3.1}$$

With the 20/20 hindsight afforded by the previous section, we can confidently answer: Yes, take $a = 6$ and $b = -1$. But the point is that in Equation (3.1) we organised this in the form of a linear relation with integer coefficients between the two numbers $\zeta(2)$ and π^2. And one can consider linear relations involving more than two numbers.

More precisely, given a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, we define an *integer relation for* $\mathbf{x}$ to be a nonzero vector $\mathbf{m} = (m_1, m_2, \ldots, m_n) \in \mathbb{Z}^n$ with integer entries such that

$$\mathbf{m} \cdot \mathbf{x} = m_1 x_1 + m_2 x_2 + \cdots + m_n x_n = 0.$$

Of course, there is no guarantee that such a magical $\mathbf{m}$ exists. (Take, for instance, $\mathbf{x} = (1, \sqrt{2})$.) This leads us to state the

Integer relation problem: Given $\mathbf{x} \in \mathbb{R}^n$, either find a "small" integer relation $\mathbf{m}$ for $\mathbf{x}$ or prove that no such "small" integer relation exists.

You may find the rather imprecise use of the adjective "small" distasteful. In that case, there is a more assertive version of the problem that fixes a bound 2^k and asks for $\mathbf{m} \in \mathbb{Z}^n$ with $\|\mathbf{m}\| \le 2^{n+k}$ or a proof that there is no $\mathbf{m} \in \mathbb{Z}^n$ with $\|\mathbf{m}\| < 2^k$.

There is yet another version that is most attractive in practice, where we take the input vector $\mathbf{x} \in \mathbb{Z}^n$ as well. The following example shows how this might come about.

Example 3.2. Let

$$\begin{aligned} x_1 &= \arctan(1) = 0.785398\ldots \\ x_2 &= \arctan(1/5) = 0.197395\ldots \\ x_3 &= \arctan(1/239) = 0.004184\ldots \end{aligned}$$

Can we find an integer relation $\mathbf{m}$ for $\mathbf{x} = (x_1, x_2, x_3)$:

$$m_1 x_1 + m_2 x_2 + m_3 x_3 = 0?$$

We turn this into a question about integers by fixing a multiplier A, say $A = 10^6$, and considering

$$m_1 \lfloor A x_1 \rfloor + m_2 \lfloor A x_2 \rfloor + m_3 \lfloor A x_3 \rfloor \approx 0,$$

that is

$$785398m_1 + 197395m_2 + 4184m_3 \approx 0.$$

There are several approaches to finding m_1, m_2, m_3 and we will be looking in more detail at one of them, the LLL algorithm. For now I will just say that this algorithm takes the matrix

$$M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 785398 & 197395 & 4184 \end{bmatrix}$$

and returns the matrix

$$\begin{bmatrix} 1 & -13 & -52 \\ -4 & 58 & 203 \\ 1 & -296 & 184 \\ 2 & 272 & 345 \end{bmatrix}$$

The first column of this matrix will be telling us that

$$785398 \cdot 1 + 197395 \cdot (-4) + 4184 \cdot 1 = 2,$$

or, going back to our original real numbers, that

$$0.785398 \cdot 1 + 0.197395 \cdot (-4) + 0.004184 \cdot 1 = \frac{2}{10^6} \approx 0.$$

In other words, $\mathbf{m} = (1, -4, 1)$ seems to be an integer relation.

Indeed, Machin found this formula in 1706:

$$\arctan(1) = 4\arctan(1/5) - \arctan(1/239).$$

It works remarkably well as a rapidly convergent approximation to $\pi = 4\arctan(1)$.

Obvious questions remain. What is this LLL algorithm? What exactly is it doing to that matrix? And how come the first column of the resulting matrix gives us the integer relation we have been looking for?

A quick answer is that LLL is a *lattice reduction algorithm*, and that the integer relation problem can be solved via lattice reduction. To make sense of this, we need to know something about lattices.

3.3 Lattice reduction

Fix a natural number n and let V be an n-dimensional real vector space endowed with an inner product $\mathbf{u} \cdot \mathbf{v}$. (You may think of V as being $\mathbb{R}^n$ with the usual dot product.)

A *lattice* in V is a subset $L \subset V$ such that there exists a basis $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ of V with

$$L = \mathrm{Span}_{\mathbb{Z}}\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\} = \{r_1\mathbf{b}_1 + r_2\mathbf{b}_2 + \cdots + r_n\mathbf{b}_n \mid r_1, r_2, \ldots, r_n \in \mathbb{Z}\}.$$

We say that $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ is a $\mathbb{Z}$-*basis* for the lattice L, and we call n the *rank* of L. The *determinant* $d(L)$ of L is the determinant of the matrix with columns $\mathbf{b}_1, \ldots, \mathbf{b}_n$. This turns out to be independent of the choice of basis.

It would be good to recall the setup of Gram–Schmidt orthogonalisation. Let $\mathbf{b}_1, \ldots, \mathbf{b}_n$ be a basis for V. For $i = 1, \ldots, n$ let $V_i = \mathrm{Span}_{\mathbb{R}}\{\mathbf{b}_1, \ldots, \mathbf{b}_i\}$.

The Gram–Schmidt process returns vectors $\mathbf{b}_1^*, \dots, \mathbf{b}_n^* \in V$ and scalars $\mu_{ij} \in \mathbb{R}$ for $1 \le j < i \le n$, defined inductively by

$$\mathbf{b}_i^* = \mathrm{proj}_{V_{i-1}^\perp}(\mathbf{b}_i) = \mathbf{b}_i - \mathrm{proj}_{V_{i-1}}(\mathbf{b}_i) = \mathbf{b}_i - \sum_{j=1}^{i-1} \mu_{ij}\mathbf{b}_j^*$$

$$\mu_{ij} = \frac{\mathbf{b}_i \cdot \mathbf{b}_j^*}{\mathbf{b}_j^* \cdot \mathbf{b}_j^*}$$

with $\mathbf{b}_1^* = \mathbf{b}_1$.

Note that, for all $i = 2, \dots, n$,

$$V_{i-1} = \mathrm{Span}_{\mathbb{R}}\{\mathbf{b}_1, \dots, \mathbf{b}_{i-1}\} = \mathrm{Span}_{\mathbb{R}}\{\mathbf{b}_1^*, \dots, \mathbf{b}_{i-1}^*\}$$

and $\mathbf{b}_1^*, \dots, \mathbf{b}_n^*$ is an orthogonal basis of V.

Let's go back to the setup of a lattice $L \subset V$. We say that a basis $\mathbf{c}_1, \dots, \mathbf{c}_n$ for L is *(LLL-)reduced* if

$$|\mu_{ij}| \le \frac{1}{2} \qquad \text{for all } 1 \le j < i \le n$$

$$\|\mathbf{c}_i^* + \mu_{i,i-1}\mathbf{c}_{i-1}^*\|^2 \ge \frac{3}{4}\|\mathbf{c}_{i-1}^*\|^2 \qquad \text{for all } 1 < i \le n.$$

Here $3/4$ can be replaced by anything in the open interval $(1/4, 1)$.

The advantage of a reduced basis is that its vectors are in some sense small:

Proposition 3.3. *If $\{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ is a reduced basis of a lattice L, then*

$$\|\mathbf{c}_j\|^2 \le 2^{i-1}\|\mathbf{c}_i^*\|^2 \qquad \textit{for all } 1 \le j \le i \le n$$

$$d(L) \le \prod_i \|\mathbf{c}_i\| \le 2^{n(n-1)/4} d(L)$$

$$\|\mathbf{c}_1\| \le 2^{(n-1)/4} d(L)^{1/n}$$

$$\|\mathbf{c}_1\|^2 \le 2^{n-1}\|\mathbf{x}\|^2 \qquad \textit{for all } \mathbf{x} \in L - \{0\}$$

It is worth dwelling a little on this last inequality. If the constant (once L is fixed) multiplier 2^{n-1} were not there, we would have that $\mathbf{c}_1$ is a shortest nonzero vector in L. This may seem like a desirable outcome (and indeed many problems rely on finding a shortest vector), but it has the great disadvantage that its time complexity is exponential in the dimension n. A reduced basis gives up some control on the size of $\mathbf{c}_1$, to the extent shown in the last inequality of the Proposition. What is gained however is that this can be computed in time polynomial in the dimension n.

The latter is achieved by the LLL reduction algorithm. Its name is derived from its authors, Arjen Lenstra, Hendrik Lenstra, and László Lovász. The algorithm starts with some given basis of L and iteratively modifies it to achieve a reduced one. It is not particularly complicated, and the exposition in the original paper [LLL82] is quite good.

Example 3.4. Going back to the situation of Example 3.2, the lattice is $\mathbb{Z}$-spanned by the columns of the matrix

$$M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 785398 & 197395 & 4184 \end{bmatrix}$$

The norms of the three basis vectors are roughly

$$785389.0, 197395.0, 4184.0$$

The LLL algorithm returns the matrix

$$\begin{bmatrix} 1 & -13 & -52 \\ -4 & 58 & 203 \\ 1 & -296 & 184 \\ 2 & 272 & 345 \end{bmatrix}$$

whose columns are the vectors in the reduced basis, with norms roughly

$$4.7, 406.4, 443.6$$

Remember that Proposition 3.3 only guarantees that the first basis vector is within a factor of $\sqrt{2^{3-1}} = 2$ of the shortest nonzero vector. In this example however, it can be checked that the first basis vector is actually a shortest vector. This does happen sometimes, an instance of the fact that the worst-case performance and the average-case performance of an algorithm can be quite different.

How does lattice reduction help with finding integer relations? (It may be good to follow along with Example 3.2.) Given $x_1, \dots, x_n \in \mathbb{R}$, let $x'_1, \dots, x'_n \in \mathbb{R}$ be close approximations (such as, for instance, truncating after a certain number of digits). Let A be a multiplier. Consider the $\mathbb{Z}$-linear map $\mathbb{Z}^n \to \mathbb{R}^{n+1}$ given by

$$\begin{bmatrix} m_1 \\ \vdots \\ m_n \end{bmatrix} \mapsto \begin{bmatrix} m_1 \\ \vdots \\ m_n \\ A\sum_i m_i x'_i \end{bmatrix}$$

Let L denote the image of this map; it is a lattice (of rank n) inside an n-dimensional subspace V of $\mathbb{R}^{n+1}$.

In other words, we consider the matrix representation of the $\mathbb{Z}$-linear map above

$$M = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ Ax'_1 & Ax'_2 & \dots & Ax'_n \end{bmatrix}$$

Letting $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$ denote the columns of this matrix, we have

$$L = \mathrm{Span}_{\mathbb{Z}}\{\mathbf{b}_1, \dots, \mathbf{b}_n\} \subset V = \mathrm{Span}_{\mathbb{R}}\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$$

The crucial observation is that

$$\begin{bmatrix} m_1 \\ \vdots \\ m_n \\ \varepsilon \end{bmatrix} \in L \qquad \text{if and only if} \qquad m_1 x'_1 + \dots + m_n x'_n = \frac{\varepsilon}{A}.$$

If ε is relatively small then (since $x'_i \approx x_i$)

$$m_1 x_1 + \dots + m_n x_n \approx 0.$$

So we are interested in finding elements of L that have small components. But, as we have seen above, the first vector in a reduced basis for L will be relatively small. So we apply LLL to the basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ of L, obtain a reduced basis $\mathbf{c}_1, \dots, \mathbf{c}_n$, and take the relation determined by the first vector $\mathbf{c}_1$.

Some answers/solutions/hints/more questions

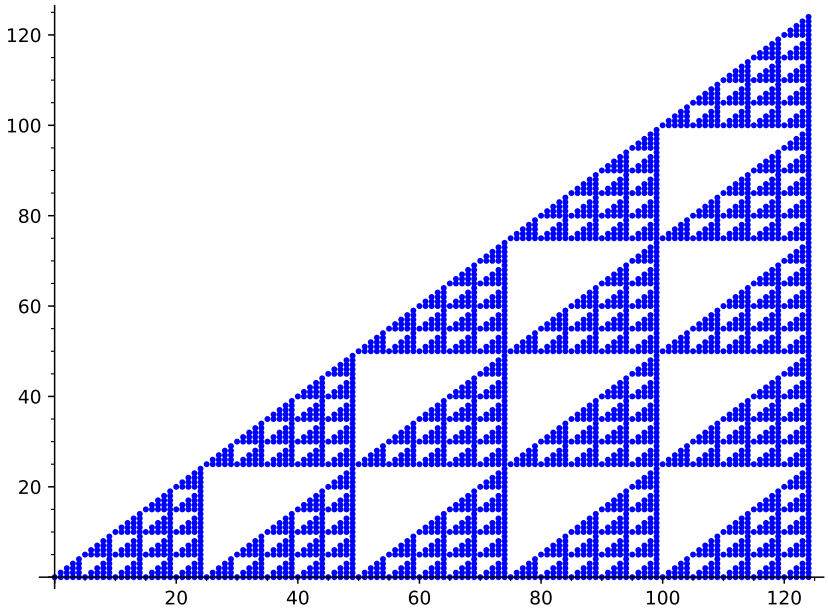
Exercise (2.2). For example:

```
sage: lst = [binomial(n, m) % 2 for n in range(2^8) for m in range(2^8)]
sage: mat = matrix(2^8, lst)
sage: p = mat.plot()
```

Exercise (2.3). Here's one approach for modulo k.

```
sage: def pascalmod(k, size=16, multicolour=True):
....:     p = list_plot([])
....:     for r in range(1, k):
....:         lstr = [(m, n) for m in range(size) for n in range(size) \
....:                 if binomial(m, n) % k == r]
....:         if multicolour:
....:             colour = hue(r/k)
....:         else:
....:             colour = 'blue'
....:         pr = list_plot(lstr, color=colour)
....:         p = p + pr
....:     return p
```

```
sage: p5 = pascalmod(5, 5^3, False)
sage: p5.show()
None
```



Exercise (3.1)**.** See [Sil13, Theorem 47.2].

Bibliography

[Cox84] David A. Cox. The arithmetic-geometric mean of Gauss. *Enseign. Math. (2)*, 30(3-4):275–330, 1984.

[Cox85] David A. Cox. Gauss and the arithmetic-geometric mean. *Notices Amer. Math. Soc.*, 32(2):147–151, 1985.

[LLL82] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász. Factoring polynomials with rational coefficients. *Math. Ann.*, 261(4):515–534, 1982.

[Sal76] Eugene Salamin. Computation of π using arithmetic-geometric mean. *Math. Comp.*, 30(135):565–570, 1976.

[Sil13] Joseph H. Silverman. *A friendly introduction to number theory*. Pearson, fourth edition, 2013.

[Sta12] Richard P. Stanley. *Enumerative combinatorics. Volume 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2012.

[Ste95] Ian Stewart. Four encounters with Sierpiński's gasket. *Math. Intelligencer*, 17(1):52–64, 1995.