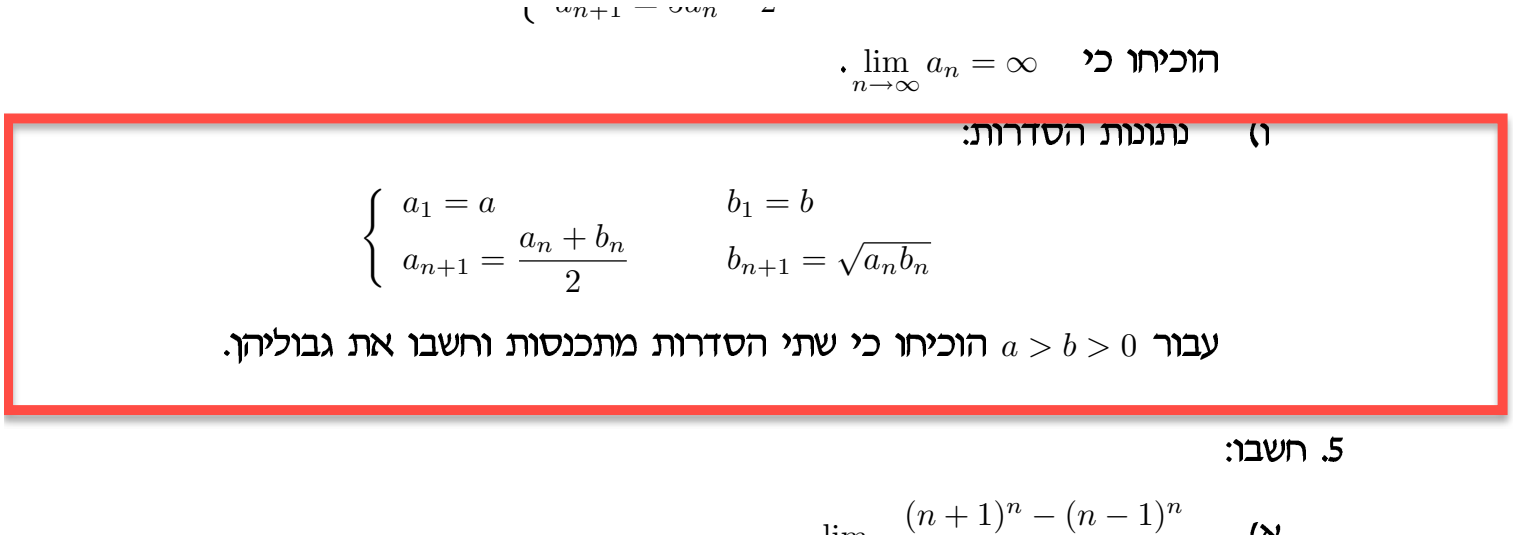
The arithmetic geometric mean (Agm)

Tamar Ziegler

Pictures by David Lehavi

This is part of exercise 5 question 4 in my calculus class Fall 2010:



הוכיחו כי $\lim_{n\to\infty} a_n = \infty$.

ו) נתונות הסדרות:

$$\begin{cases} a_1 = a & b_1 = b \\ a_{n+1} = \dfrac{a_n + b_n}{2} & b_{n+1} = \sqrt{a_n b_n} \end{cases}$$

עבור $a > b > 0$ הוכיחו כי שתי הסדרות מתכנסות וחשבו את גבוליהן.

5. חשבו:

$$\lim \quad (n+1)^n - (n-1)^n$$

Here is the translation (and index change):

$$a_0 = a \qquad b_0 = b$$

$$a_{n+1} = \frac{a_n + b_n}{2} \qquad b_{n+1} = \sqrt{a_n b_n} \qquad n = 0, 1, \ldots$$

For $a > b > 0$ show that both sequences converge and calculate their limits.

Why do the limits exist ?

Why do the limits exist ?

By the mean inequality:

$$\frac{a+b}{2} \geq \sqrt{ab}$$

Why do the limits exist ?

By the mean inequality:

$$\frac{a+b}{2} \geq \sqrt{ab}$$

Therefore

$$a_1 \geq b_1$$

and by induction

$$a_n \geq b_n.$$

Why do the limits exist ?

By the mean inequality:

$$\frac{a+b}{2} \geq \sqrt{ab}$$

Therefore

$$a_1 \geq b_1$$

and by induction

$$a_n \geq b_n.$$

Also

$$a_n \geq \frac{a_n + b_n}{2} = a_{n+1} \geq b_{n+1} = \sqrt{a_n b_n} \geq b_n$$

We have

$$a \geq a_1 \geq \ldots \geq a_n \geq a_{n+1} \geq b_{n+1} \geq b_n \geq \ldots \geq b_1 \geq b$$

We have

$$a \geq a_1 \geq \ldots \geq a_n \geq a_{n+1} \geq b_{n+1} \geq b_n \geq \ldots \geq b_1 \geq b$$

The sequence $\{a_n\}$ is decreasing and bounded below by b, and the sequence $\{b_n\}$ is increasing and bounded above by a. So both sequences converge !

We have

$$a \geq a_1 \geq \ldots \geq a_n \geq a_{n+1} \geq b_{n+1} \geq b_n \geq \ldots \geq b_1 \geq b$$

The sequence $\{a_n\}$ is decreasing and bounded below by b, and the sequence $\{b_n\}$ is increasing and bounded above by a. So both sequences converge !

Actually, they converge to a common limit:

$$0 \leq a_{n+1} - b_{n+1} \leq a_{n+1} - b_n = \frac{a_n + b_n}{2} - b_n = \frac{a_n - b_n}{2}$$

We have

$$a \geq a_1 \geq \ldots \geq a_n \geq a_{n+1} \geq b_{n+1} \geq b_n \geq \ldots \geq b_1 \geq b$$

The sequence $\{a_n\}$ is decreasing and bounded below by b, and the sequence $\{b_n\}$ is increasing and bounded above by a. So both sequences converge !

Actually, they converge to a common limit:

$$0 \leq a_{n+1} - b_{n+1} \leq a_{n+1} - b_n = \frac{a_n + b_n}{2} - b_n = \frac{a_n - b_n}{2}$$

Inductively:

$$0 \leq a_n - b_n \leq \frac{a - b}{2^n}$$

Define $M(a,b)$ - the *arithmetic geometric mean of* a,b to be the common limit of the sequences $\{a_n\},\{b_n\}$

Arithmetic geometric mean

$$M(a,b)=\lim_{n\to\infty} a_n=\lim_{n\to\infty} b_n$$

Define $M(a,b)$ - the *arithmetic geometric mean of* a,b to be the common limit of the sequences $\{a_n\},\{b_n\}$

Arithmetic geometric mean

$$M(a,b)=\lim_{n\to\infty} a_n=\lim_{n\to\infty} b_n$$

Simple observations:

$$M(a,b)=M(a_1,b_1)=M(a_2,b_2)=\dots$$

Define $M(a,b)$ - the *arithmetic geometric mean of* a,b to be the common limit of the sequences $\{a_n\},\{b_n\}$

Arithmetic geometric mean

$$M(a,b)=\lim_{n\to\infty} a_n=\lim_{n\to\infty} b_n$$

Simple observations:

$$M(a,b)=M(a_1,b_1)=M(a_2,b_2)=\dots$$

$$M(\lambda a,\lambda b)=\lambda M(a,b).$$

Define $M(a,b)$ - the *arithmetic geometric mean of* a,b to be the common limit of the sequences $\{a_n\},\{b_n\}$

Arithmetic geometric mean

$$M(a,b)=\lim_{n\to\infty} a_n=\lim_{n\to\infty} b_n$$

Simple observations:

$$M(a,b)=M(a_1,b_1)=M(a_2,b_2)=\dots$$

$$M(\lambda a,\lambda b)=\lambda M(a,b).$$

$$M(a,a)=a$$

Ok then, what about calculating the limit ? Hmm...

$$M(\sqrt{2}, 1) = 1.19814023473559220 74\ldots$$

(accuracy to 19 decimal places).

$$M(\sqrt{2}, 1) = 1.19814023473559220 74\ldots$$

(accuracy to 19 decimal places).

All entries of the following table are rounded to 21 decimal places.

n	a_n	b_n
0	1.414213562373905048802	1.000000000000000000000

$$M(\sqrt{2},1) = 1.19814023473559220 74\ldots$$

(accuracy to 19 decimal places).

All entries of the following table are rounded to 21 decimal places.

n	a_n	b_n
0	1.414213562373905048802	1.000000000000000000000
1	1.207106781186547524401	1.189207115002721066717

$$M(\sqrt{2},1) = 1.19814023473559220 74\ldots$$

(accuracy to 19 decimal places).

All entries of the following table are rounded to 21 decimal places.

n	a_n	b_n
0	1.414213562373905048802	1.000000000000000000000
1	1.207106781186547524401	1.189207115002721066717
2	1.198156948094634295559	1.198123521493120122607

$$M(\sqrt{2},1) = 1.19814023473559220 74\ldots$$

(accuracy to 19 decimal places).

All entries of the following table are rounded to 21 decimal places.

n	a_n	b_n
0	1.414213562373905048802	1.000000000000000000000
1	1.207106781186547524401	1.189207115002721066717
2	1.198156948094634295559	1.198123521493120122607
3	1.198140234793877209083	1.198140234677307205798

$$M(\sqrt{2},1) = 1.19814023473559220 74\ldots$$

(accuracy to 19 decimal places).

All entries of the following table are rounded to 21 decimal places.

n	a_n	b_n
0	1.414213562373905048802	1.000000000000000000000
1	1.207106781186547524401	1.189207115002721066717
2	1.198156948094634295559	1.198123521493120122607
3	1.198140234793877209083	1.198140234677307205798
4	1.198140234735592207441	1.198140234735592207439

$$M(\sqrt{2},1) = 1.19814023473559220 74\ldots$$

(accuracy to 19 decimal places).

All entries of the following table are rounded to 21 decimal places.

n	a_n	b_n
0	1.414213562373905048802	1.000000000000000000000
1	1.207106781186547524401	1.189207115002721066717
2	1.198156948094634295559	1.198123521493120122607
3	1.198140234793877209083	1.198140234677307205798
4	1.198140234735592207441	1.198140234735592207439

Gauss calculated these himself !!! (This table appears in his manuscript from 1800).

$$M(\sqrt{2},1) = 1.19814023473559220 74\ldots$$

(accuracy to 19 decimal places).

All entries of the following table are rounded to 21 decimal places.

n	a_n	b_n
0	1.414213562373905048802	1.000000000000000000000
1	1.207106781186547524401	1.189207115002721066717
2	1.198156948094634295559	1.198123521493120122607
3	1.198140234793877209083	1.198140234677307205798
4	1.198140234735592207441	1.198140234735592207439

Gauss calculated these himself !!! (This table appears in his manuscript from 1800).

After 4 (!!!) iterations we get 19 digits accuracy, the convergence is really really (really) fast. Much better than $\frac{b-a}{2^n} \sim \frac{2}{100}$.

Better calculation:

$$\delta_{n+1} = a_{n+1} - b_{n+1} = \frac{a_n + b_n}{2} - \sqrt{a_n b_n}$$

Better calculation:

$$\delta_{n+1} = a_{n+1} - b_{n+1} = \frac{a_n + b_n}{2} - \sqrt{a_n b_n}$$
$$= \frac{1}{2}(a_n + b_n - 2\sqrt{a_n b_n}) = \frac{1}{2}(\sqrt{a_n} - \sqrt{b_n})^2$$

Better calculation:

$$\delta_{n+1} = a_{n+1} - b_{n+1} = \frac{a_n + b_n}{2} - \sqrt{a_n b_n}$$
$$= \frac{1}{2}(a_n + b_n - 2\sqrt{a_n b_n}) = \frac{1}{2}(\sqrt{a_n} - \sqrt{b_n})^2$$

On the other hand

$$\delta_n = a_n - b_n = (\sqrt{a_n} - \sqrt{b_n})(\sqrt{a_n} + \sqrt{b_n})$$

Better calculation:

$$\delta_{n+1} = a_{n+1} - b_{n+1} = \frac{a_n + b_n}{2} - \sqrt{a_n b_n}$$
$$= \frac{1}{2}(a_n + b_n - 2\sqrt{a_n b_n}) = \frac{1}{2}(\sqrt{a_n} - \sqrt{b_n})^2$$

On the other hand

$$\delta_n = a_n - b_n = (\sqrt{a_n} - \sqrt{b_n})(\sqrt{a_n} + \sqrt{b_n})$$

So

$$\frac{\delta_{n+1}}{\delta_n^2} = \frac{1}{2(\sqrt{a_n} + \sqrt{b_n})^2} \xrightarrow[n\to\infty]{} \frac{1}{8M(a,b)}$$

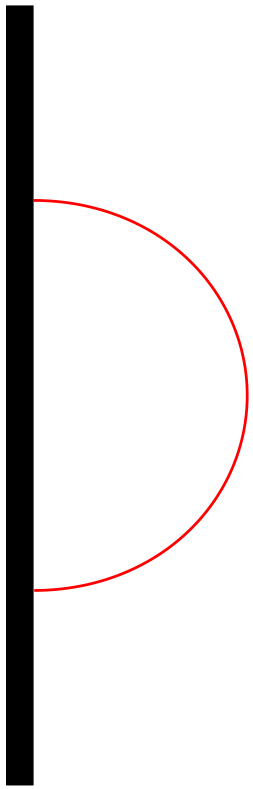
convergence rate

$$\delta_{n+1} \sim C\delta_n^2$$

Indeed, convergence rate is much better - it is quadratic: if at stage n we have k digits accuracy, then at stage $n+1$ our accuracy is $2k$ digits.

The story of the lemniscate

In 1691 - Jacob Bernoulli was working on the following problem:
A thin elastic rod is bent until the two ends are perpendicular to a given line:



Bernoulli showed that the upper half of the curve is given by an equation of the form

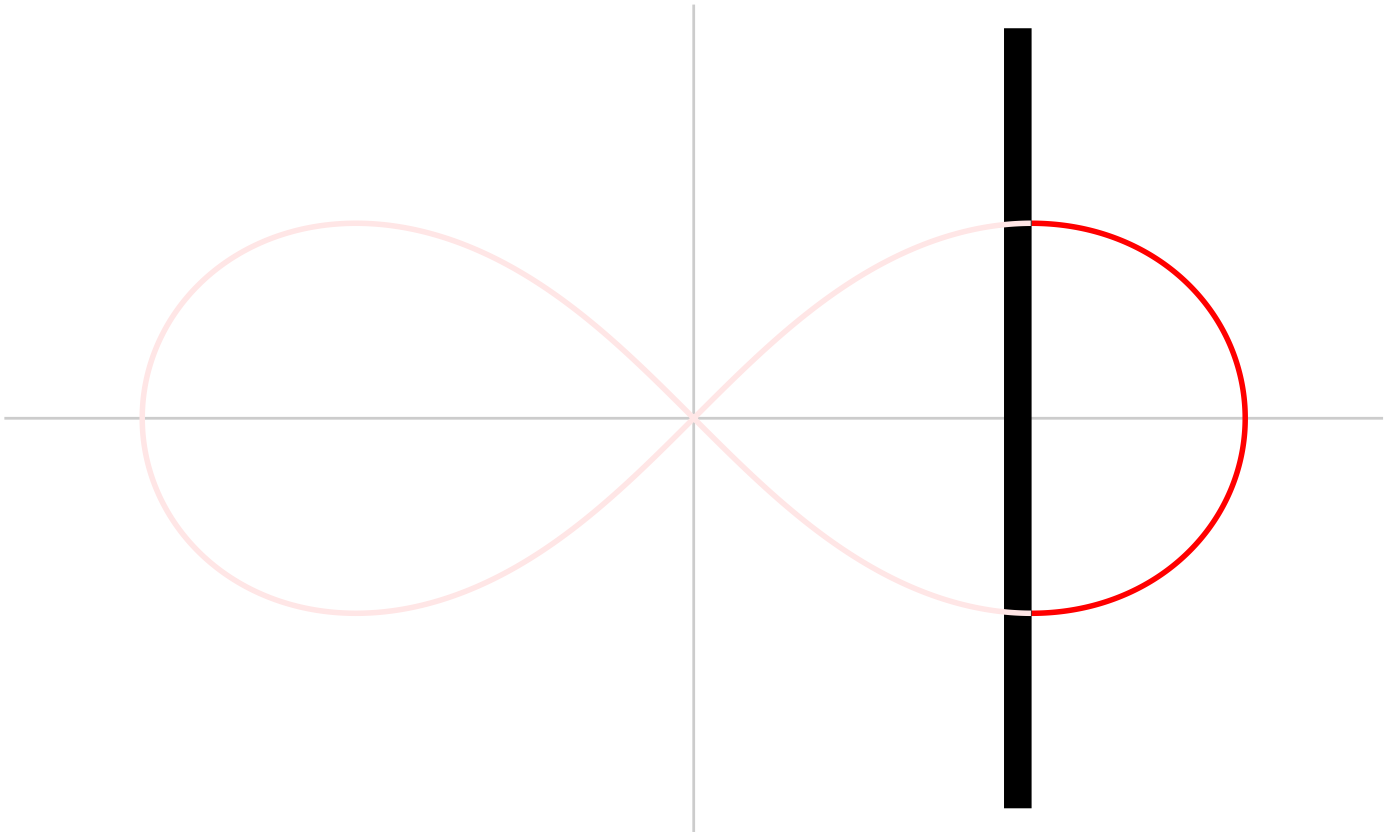
$$y = \int_0^x \frac{z^2}{\sqrt{a^4 - z^4}} dz$$

How is this related to the lemniscate?

Bernoulli sought 'an algebraic curve whose rectification should agree with the rectification of the elastic curve'

In 1694 he found the *lemniscate* (Greek for ribbon):

$$x^2 + y^2 = a\sqrt{x^2 - y^2}$$



1694 - Jacob's younger brother, Johann independently discovered the lemniscate !

Both papers appeared in Acta Eruditorum:
Jacob - September 1694
Johann -October 1694

Anyway, how is this related to agm ???

Stirling (1730):

The length of $1/4$ lemniscate:

$$A = \int_0^1 \frac{1}{\sqrt{1-z^4}} dz = 1.21102877714605987$$

and also:

$$B = \int_0^1 \frac{z^2}{\sqrt{1-z^4}} dz = .5990701173677961 1$$

Observe that

$$2 \cdot B = 1.1981402347355922 2$$

Stirling (1730):

The length of $1/4$ lemniscate:

$$A = \int_0^1 \frac{1}{\sqrt{1-z^4}} dz = 1.21102877714605987$$

and also:

$$B = \int_0^1 \frac{z^2}{\sqrt{1-z^4}} dz = .5990701173677961 1$$

Observe that

$$2 \cdot B = 1.1981402347355922 2$$

Which agrees with $M(\sqrt{2}, 1)$ to 16 decimal places !

Carl Freidrich Gauss

$$M(\sqrt{2},1) = 2 \cdot B = \int_0^1 \frac{z^2}{\sqrt{1-z^4}} dz$$

In 1786 Euler showed

$$A \cdot B = \int_0^1 \frac{1}{\sqrt{1-z^4}} dz \cdot \int_0^1 \frac{z^2}{\sqrt{1-z^4}} dz = \frac{\pi}{4}$$

Coupled with Eulers result we get

$$\tilde{\omega} = 2\int_0^1 \frac{1}{\sqrt{1-z^4}}dz = \frac{\pi}{M(\sqrt{2},1)} = \text{length of } 1/2 \text{ Lemniscate}$$

This is a special notation of Gauss.

Coupled with Eulers result we get

$$\tilde{\omega} = 2\int_0^1 \frac{1}{\sqrt{1-z^4}}dz = \frac{\pi}{M(\sqrt{2},1)} = \text{length of } 1/2 \text{ Lemniscate}$$

This is a special notation of Gauss.

Compare with:

$$2\int_0^1 \frac{1}{\sqrt{1-z^2}}dz = \pi = \text{length of } 1/2 \text{ circle}$$

He gave a special name to the function: arclmnsin x

$$\text{arclmnsin } x = \int_0^x \frac{1}{\sqrt{1-z^4}} dz$$

Then $2 \cdot \text{arclmnsin} 1 = \tilde{\omega} = \frac{\pi}{M(\sqrt{2},1)}$ - the period of the lemniscate.

Compare to

$$\arcsin x = \int_0^x \frac{1}{\sqrt{1-z^2}} dz$$

with $2 \cdot \arcsin 1 = \pi$ - the period of the circle.

Theorem [Gauss]

$$M(a,b)\cdot\int_0^{\frac{\pi}{2}}\frac{1}{\sqrt{a^2(\cos\phi)^2+b^2(\sin\phi)^2}}d\phi=\frac{\pi}{2}$$

(exercise: write the equation of the lemniscare in polar coordinates, then write down the expression for its arclength)

Theorem [Gauss]

$$M(a,b)\cdot\int_0^{\frac{\pi}{2}}\frac{1}{\sqrt{a^2(\cos\phi)^2+b^2(\sin\phi)^2}}d\phi=\frac{\pi}{2}$$

(exercise: write the equation of the lemniscare in polar coordinates, then write down the expression for its arclength)

Proof: Denote

$$I(a,b)=\int_0^{\frac{\pi}{2}}\frac{1}{\sqrt{a^2(\cos\phi)^2+b^2(\sin\phi)^2}}d\phi$$

Key step: Show that

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

where

$$a_1=\frac{a+b}{2}\qquad b_1=\sqrt{ab}$$

If we do that, we are done, for then

$$I(a,b) = I(a_1,b_1) = I(a_2,b_2) = \dots$$

Therefore

$$I(a,b) = \lim_{n\to\infty} I(a_n,b_n) = \lim \int_0^{\frac{\pi}{2}} \frac{1}{\sqrt{a_n^2(\cos\phi)^2 + b_n^2(\sin\phi)^2}} d\phi$$

The integrand converges uniformly in ϕ to

$$\int_0^{\frac{\pi}{2}} \frac{1}{\sqrt{M(a,b)^2(\cos\phi)^2 + M(a,b)^2(\sin\phi)^2}} d\phi = \frac{1}{M(a,b)} \cdot \frac{\pi}{2}$$

How do we show the key step $I(a,b) = I(a_1.b_1)$?

How do we show the key step $I(a,b) = I(a_1.b_1)$?

Ingenious change of variables !

$$\sin\phi = \frac{2a\sin\phi'}{a+b+(a-b)(\sin\phi')^2}$$

The range $0 \leq \phi' \leq \frac{\pi}{2}$ corresponds to $0 \leq \phi \leq \frac{\pi}{2}$.

How do we show the key step $I(a,b) = I(a_1.b_1)$?

Ingenious change of variables !

$$\sin\phi = \frac{2a\sin\phi'}{a+b+(a-b)(\sin\phi')^2}$$

The range $0 \leq \phi' \leq \frac{\pi}{2}$ corresponds to $0 \leq \phi \leq \frac{\pi}{2}$.

Gauss then writes 'after the development has been made correctly, it will be seen that'

$$\frac{1}{\sqrt{a^2(\cos\phi)^2 + b^2(\sin\phi)^2}}d\phi = \frac{1}{\sqrt{a_1^2(\cos\phi')^2 + b_1^2(\sin\phi')^2}}d\phi'$$

(Challenge - try to fill in the details ...)

How the **** did Gauss come up with this substitution ???

Lets go back to the integral we are trying to compute:

$$2\int_0^1 \frac{1}{\sqrt{1-z^4}}dz$$

Why is it hard to compute?

How the **** did Gauss come up with this substitution ???

Lets go back to the integral we are trying to compute:

$$2\int_0^1 \frac{1}{\sqrt{1-z^4}}dz$$

Why is it hard to compute?

Try all you like, the integral

$$2\int_0^x \frac{1}{\sqrt{1-z^4}}dz$$

can not be expressed using elementary functions: polynomials, $\sin, \cos, \exp$ and their inverses.

Compare to

$$2\int_0^x \frac{1}{\sqrt{1-z^2}}dz$$

Compare to

$$2\int_0^x \frac{1}{\sqrt{1-z^2}}dz$$

After a change of variables $z = \sin t$ we get

$$2\int_0^{\arcsin x} \frac{\cos t}{\cos t}dt = 2\int_0^{\arcsin x} 1dt = 2\arcsin x$$

For $x = 1$ we get a period of the function $\sin x$ - namely π.

In fact we can calculate all integral of the form

$$\int \frac{P(\sin t, \cos t)}{Q(\sin t, \cos t)} dt$$

for any polynomials in two variables P, Q like

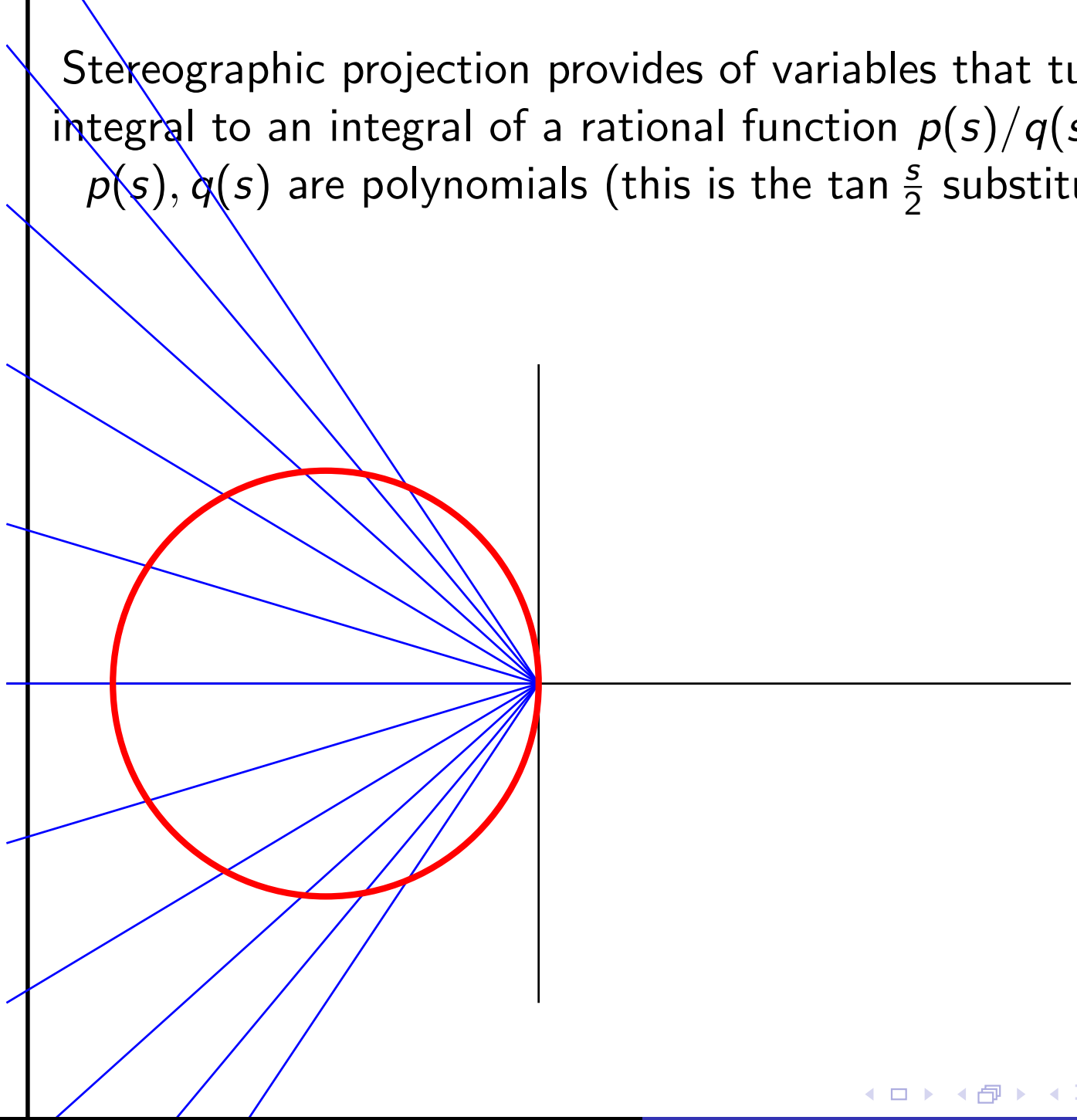
$$\int 1 dt \qquad \text{which is what we got before}$$

but also

$$\int \frac{(\sin t)^2(\cos t)^5 + \sin t}{(\cos t)^{16} + (\sin t)^7(\cos t)^{13}} dt$$

HOW SO?

Stereographic projection provides of variables that turns the integral to an integral of a rational function $p(s)/q(s)$ where $p(s), q(s)$ are polynomials (this is the tan $\frac{s}{2}$ substitution).



Back to

$$\int_0^1 \frac{1}{\sqrt{1-z^4}}dz = \int_0^1 \frac{1}{\sqrt{(1+z^2)(1-z)(1+z)}}dz$$

Back to

$$\int_0^1 \frac{1}{\sqrt{1-z^4}}dz = \int_0^1 \frac{1}{\sqrt{(1+z^2)(1-z)(1+z)}}dz$$

Making a change of variable $z = 1 - 1/x$ we get

$$\int_{-\infty}^0 \frac{1}{\sqrt{(x^2-2x+2)(2x-1)}}dx$$

Back to

$$\int_0^1 \frac{1}{\sqrt{1-z^4}}dz = \int_0^1 \frac{1}{\sqrt{(1+z^2)(1-z)(1+z)}}dz$$

Making a change of variable $z = 1 - 1/x$ we get

$$\int_{-\infty}^0 \frac{1}{\sqrt{(x^2-2x+2)(2x-1)}}dx$$

Consider the equation

$$y^2 = (x^2-2x+2)(2x-1)$$

Back to

$$\int_0^1 \frac{1}{\sqrt{1-z^4}}dz = \int_0^1 \frac{1}{\sqrt{(1+z^2)(1-z)(1+z)}}dz$$

Making a change of variable $z = 1 - 1/x$ we get

$$\int_{-\infty}^0 \frac{1}{\sqrt{(x^2-2x+2)(2x-1)}}dx$$

Consider the equation

$$y^2 = (x^2-2x+2)(2x-1)$$

This is a special case of

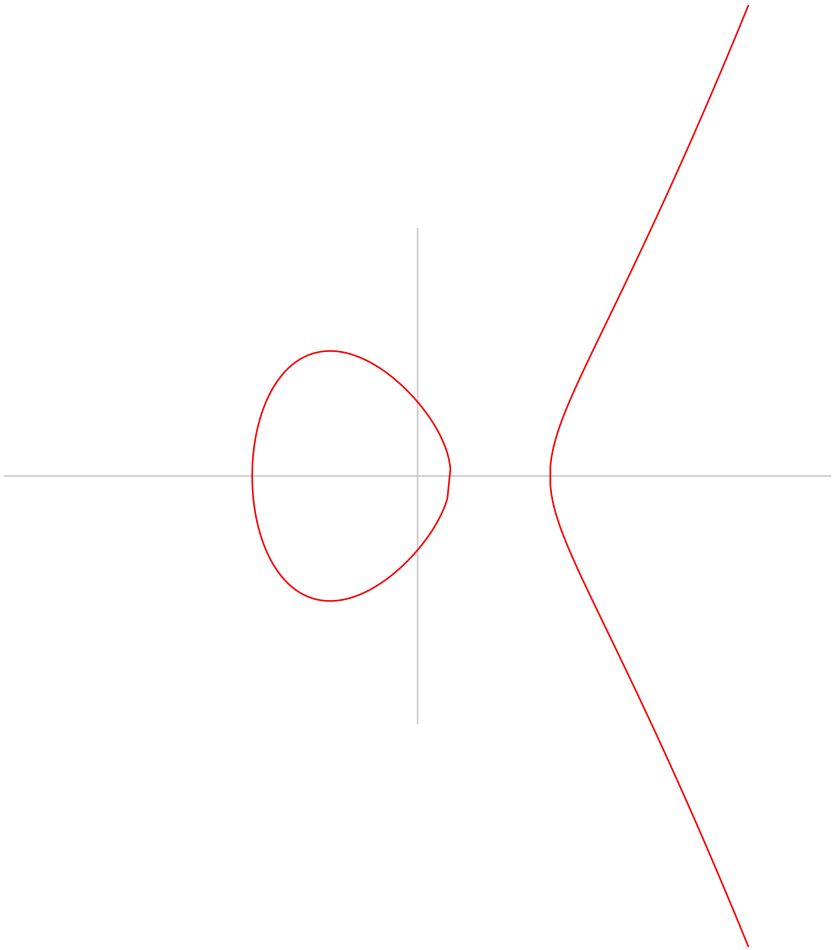
$$y^2 = ax^3 + bx^2 + cx + d$$

The graph describing the solutions to an equation of this type called an *elliptic curve*, and an integral

$$\int \frac{dx}{y(x)}$$

is called an *elliptic integral* (of the first kind).

If all the roots of P are real, then the curve looks like this.



There is no good 'stereographic projection' here (we can only get a map that is $2:1$)

Equations behave much better when we look at them over the complex numbers $\mathbb{C}$.

Recall our equation was

$$y^2 = ax^3 + bx^2 + cx + d$$

and we now think of y, x as taking complex values.

We take one more step - we *projectivize*. We first make the equation homogeneous:

$$\left(\frac{y}{z}\right)^2 = a\left(\frac{x}{z}\right)^3 + b\left(\frac{x}{z}\right)^2 + c\left(\frac{x}{z}\right) + d$$

We take one more step - we *projectivize*. We first make the equation homogeneous:

$$\left(\frac{y}{z}\right)^2 = a\left(\frac{x}{z}\right)^3 + b\left(\frac{x}{z}\right)^2 + c\left(\frac{x}{z}\right) + d$$

multiplying by z^3 we get

$$y^2z = ax^3 + bx^2z + cxz^2 + dz^3$$

(If we substitute $z = 1$ we get our original equation).

We now look at all triples (x, y, z) satisfying the equation,

$$y^2z = ax^3 + bx^2z + cxz^2 + dz^3$$

but we identify two triples

$$(x, y, z) \sim \lambda(x, y, z) \qquad \lambda \in \mathbb{C}, \lambda \neq 0$$

We now look at all triples (x, y, z) satisfying the equation,

$$y^2z = ax^3 + bx^2z + cxz^2 + dz^3$$

but we identify two triples

$$(x, y, z) \sim \lambda(x, y, z) \qquad \lambda \in \mathbb{C}, \lambda \neq 0$$

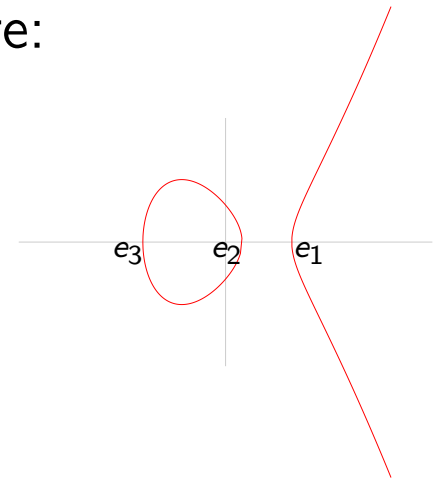
The set of all triple $(x, y, z) \in \mathbb{C}^3$ with this identification is denoted $\mathbb{PC}^3$ (or $\mathbb{P}^2\mathbb{C}$). This is a way to make the set of solutions compact.

DON'T GET LOST JUST YET !

Now we can allow ourselves to take linear transformations of $\mathbb{PC}^3$ (möbius transformations), and actually assume that our elliptic curve has 3 real roots $e_1 \geq e_2 \geq e_3$ and we are looking at the integral

$$\int_{e_3}^{e_2} \frac{1}{\sqrt{(x-e_1)(x-e_2)(x-e_3)}} dx$$

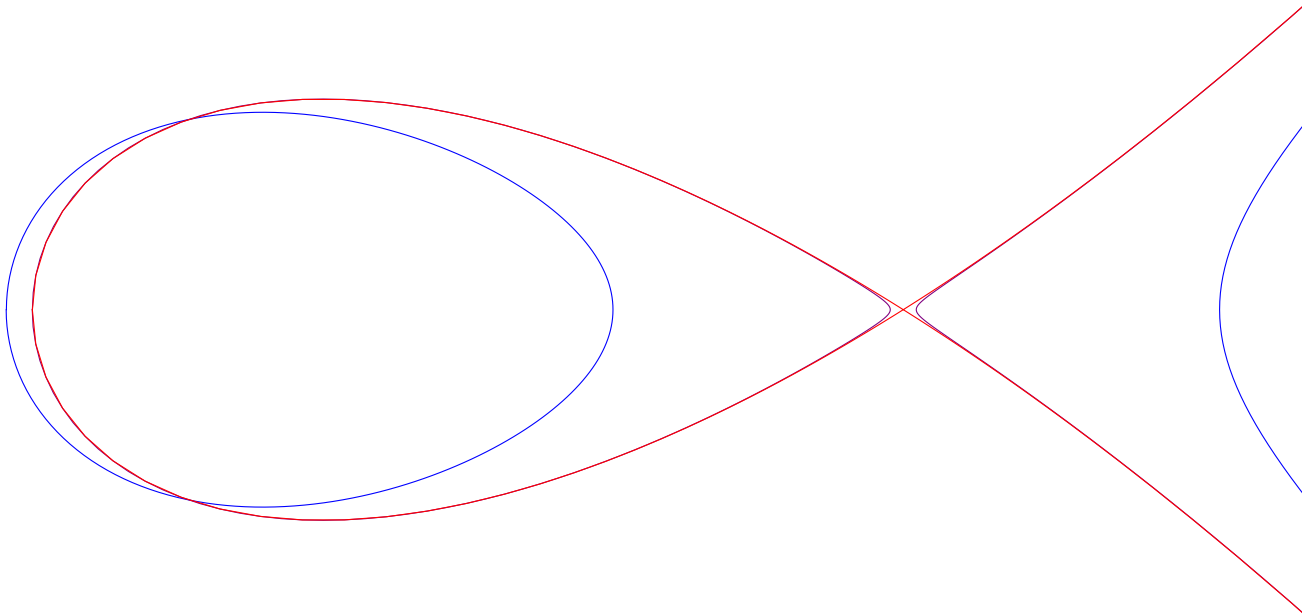
Remember the picture:



At last we are ready to describe what the agm substitution does:

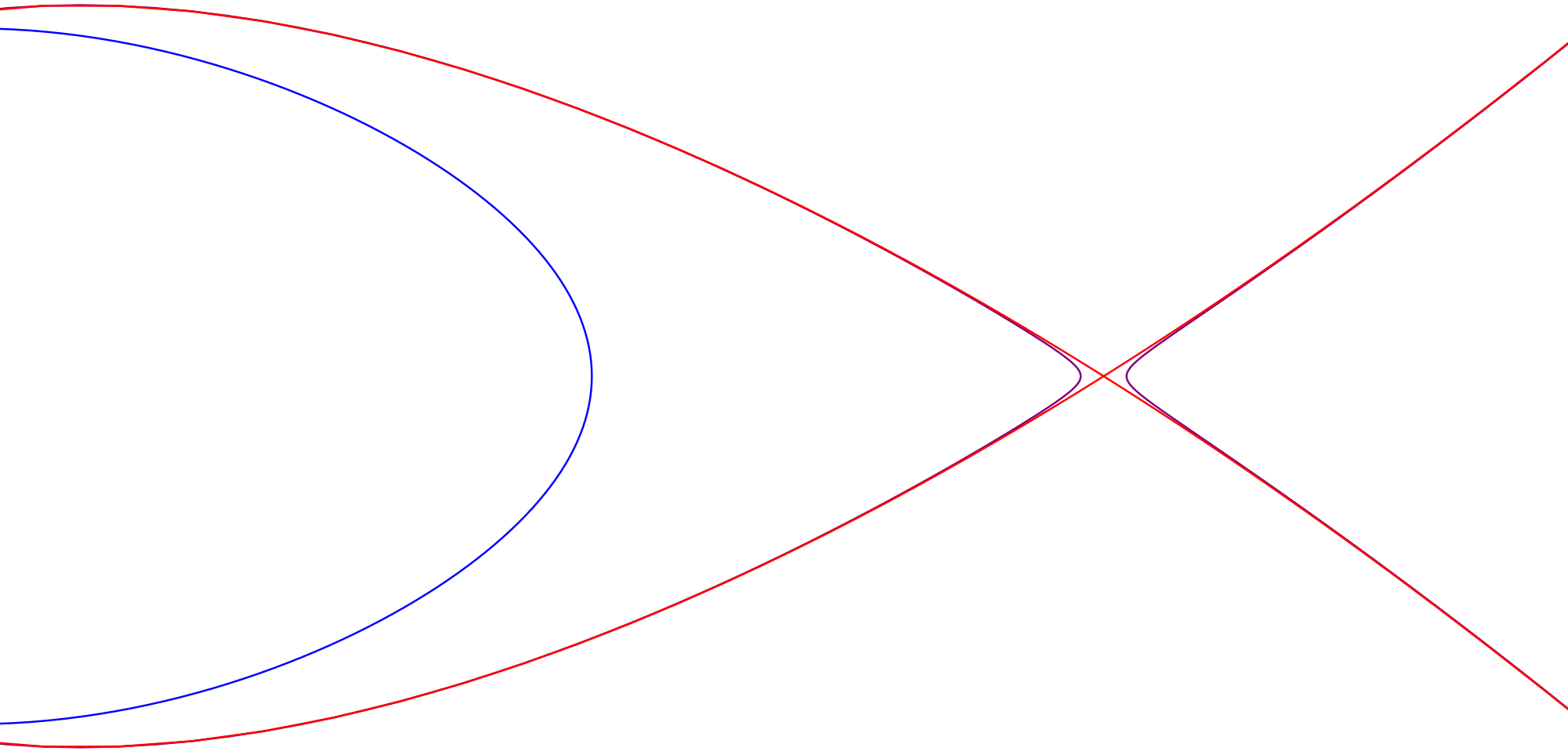
Each time we apply the ingenious substitution of Gauss, we change the elliptic curve in the in integral. What happens to the geometric picture?

Each time we apply the ingenious substitution of Gauss, we change the elliptic curve in the in integral. What happens to the geometric picture?

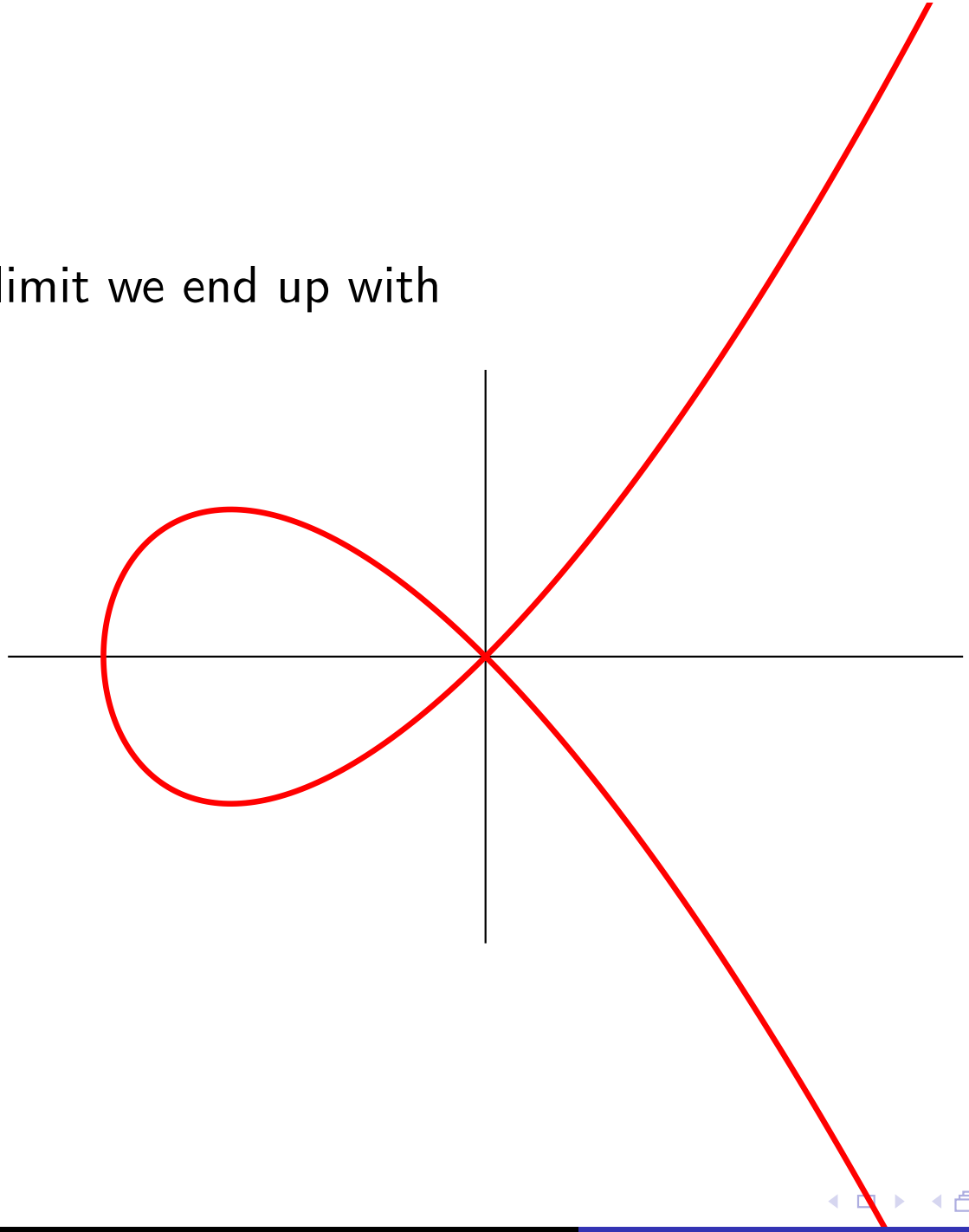


This is a picture of 2 iterations (recall that 2 iterations of the agm give us 4 decimal point precision).

Lets zoom in: e_1, e_2 are getting really close ...



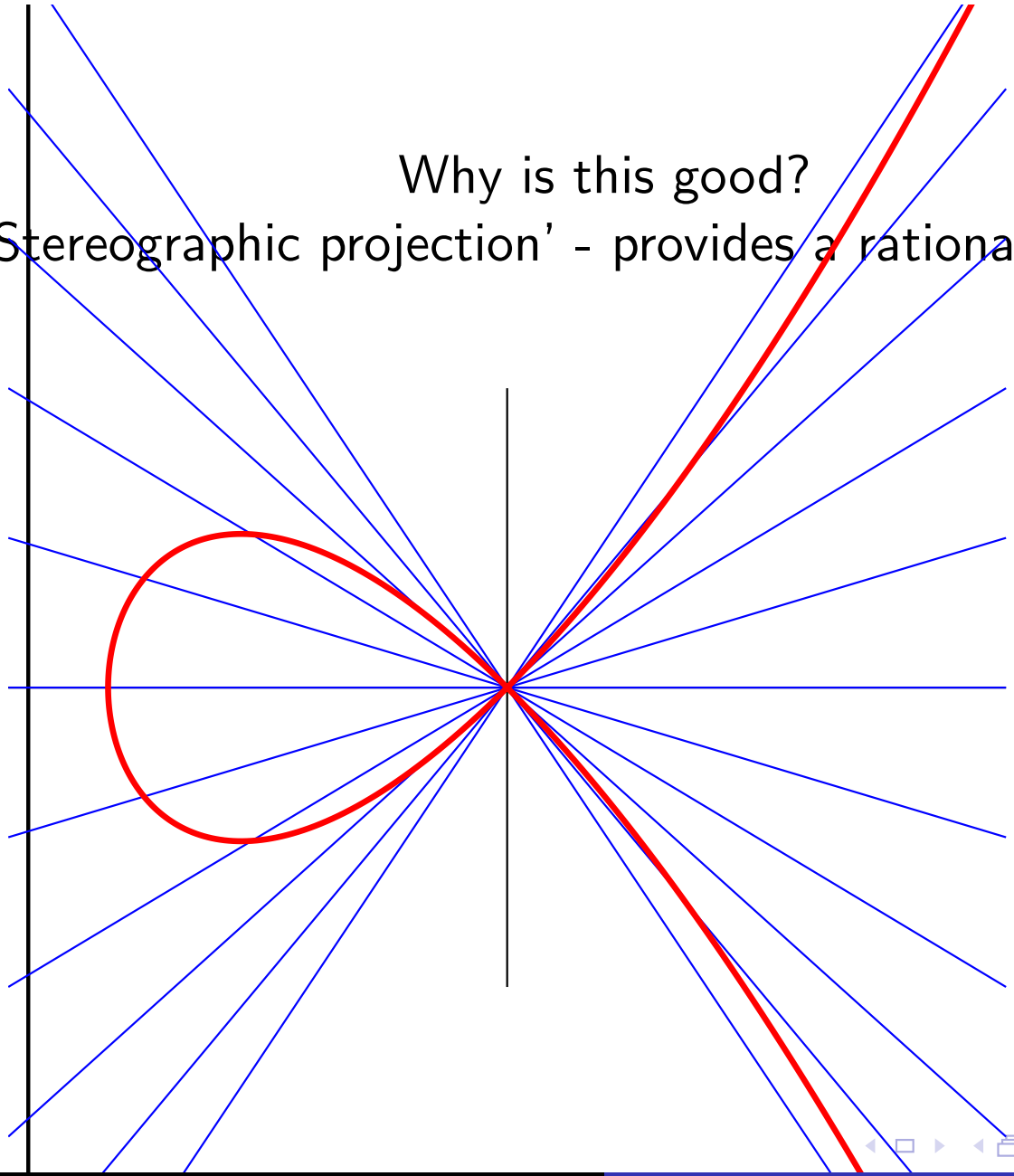
In the limit we end up with



Why is this good?

Why is this good?

'Stereographic projection' - provides a rationalization!!!



At the integral level - this is providing a sequence of approximations ending with something of the form

$$\int_{e_3'}^{e_2'} \frac{1}{\sqrt{(x-e_2')^2(x-e_3')}} dx$$

(these are not the original e_2, e_3 we stated out with)

HOW?

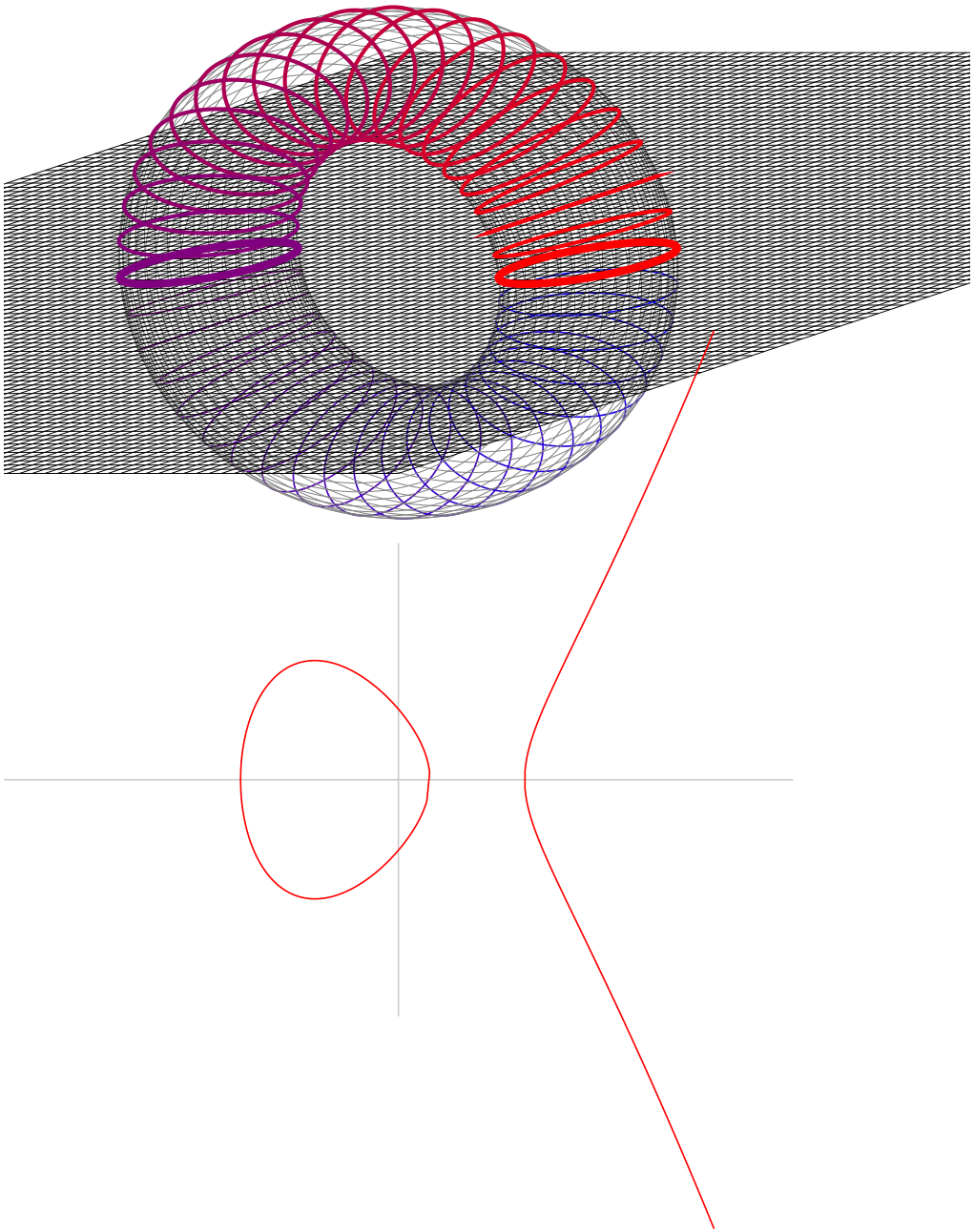
It turns out that a miracle happens and the set of solutions to the equation

$$y^2z = ax^3 + bx^2z + cxz^2 + dz^3$$

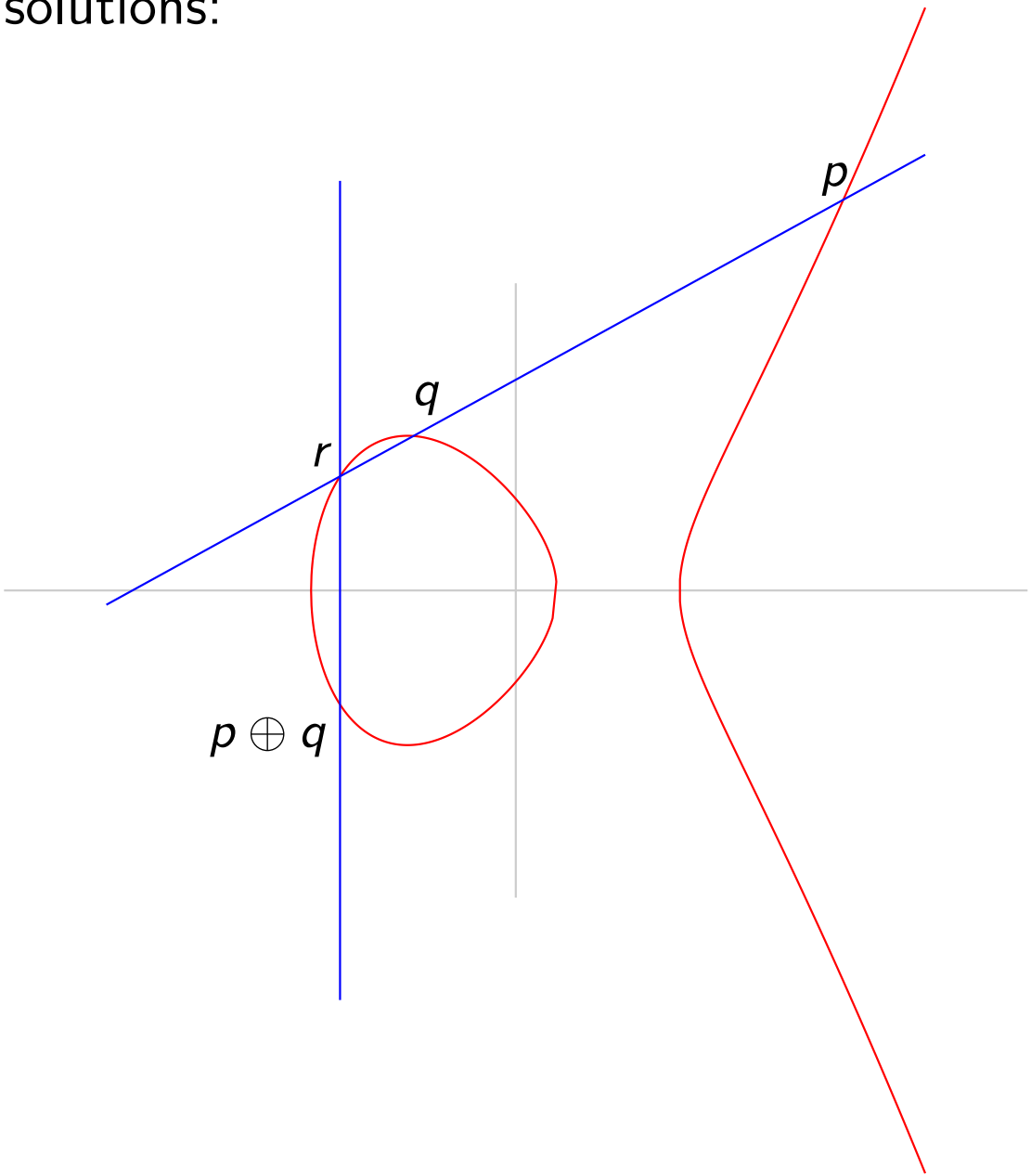
form a group, and this group is isomorphic to $\mathbb{C}/\Lambda$ where

$$\Lambda = \mathbb{Z} + \omega\mathbb{Z}$$

is a lattice.



Here is what the group operation looks like on the subset of real solutions:



Where were we ...

$$F : \mathbb{C}/(\mathbb{Z} + \omega\mathbb{Z}) \hookrightarrow \mathbb{P}\mathbb{C}^3$$
$$z \longmapsto [\wp(z) : \wp'(z) : 1]$$
$$0 \longmapsto [0 : 1 : 0]$$

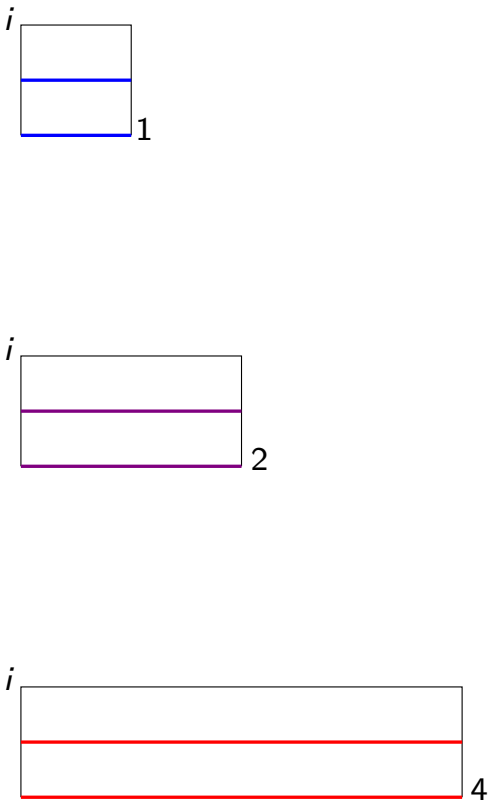
The image of F is an elliptic curve E. (The map from $\mathbb{C}/(\mathbb{Z} + \omega\mathbb{Z})$ onto its image is a homomorphism of (abelian) groups).

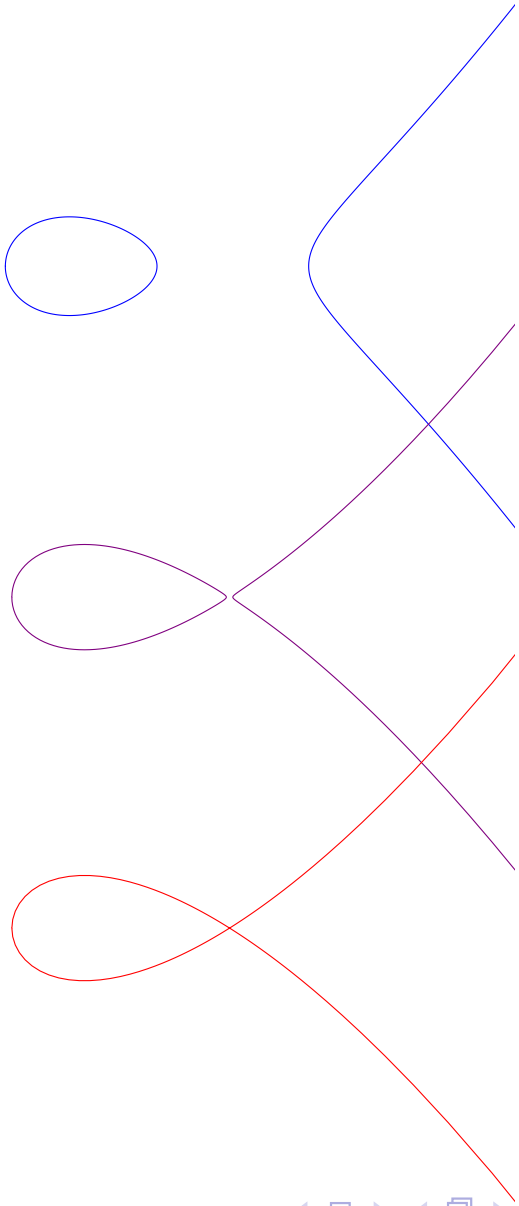
$$x = \wp(z); \quad dx = \wp'(z)dz = ydz$$

Thus

$$\int_{e_3}^{e_2} \frac{dx}{y} = \int_{F^{-1}([e_3,e_2])} 1dz$$

What happens with agm? division of fundamental domain by 2 !





Who is $\wp$? This is the Weierstrass function, it is defined as follows:

Weierstrass $\wp$ function

$$\wp(z,\Lambda) = \frac{1}{z^2} + \sum_{w\in\Lambda} \frac{1}{(z-w)^2} - \frac{1}{w^2}$$

Recall the map

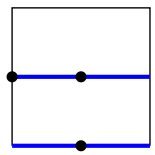
$$\begin{aligned} F : \mathbb{C}/\Lambda &\hookrightarrow \mathbb{P}\mathbb{C}^3 \\ z &\longmapsto [\wp(z,\Lambda) : \wp'(z,\Lambda) : 1] \end{aligned}$$

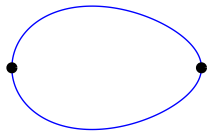
Example 0

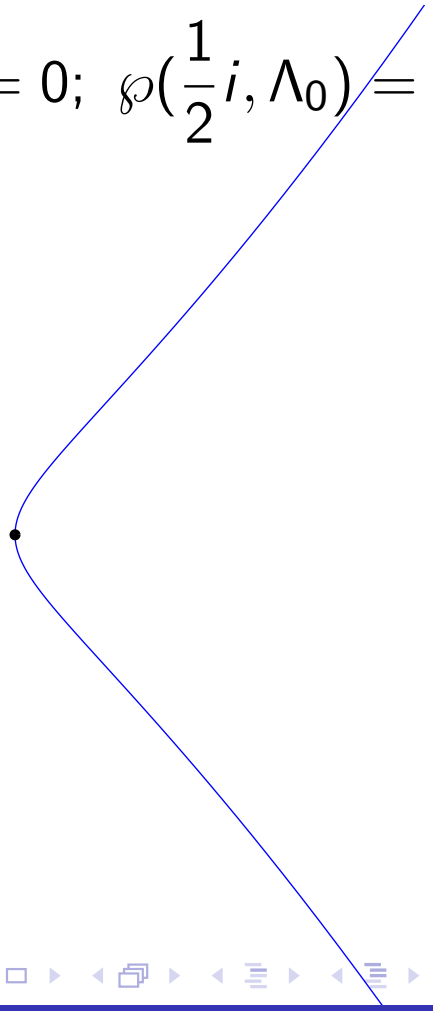
This is the zero iteration of agm:

$$\Lambda_0 = \mathbb{Z} + i\mathbb{Z} \quad \leftrightarrow \quad y_0 = (x-e)x(x+e)$$

$$\wp(0,\Lambda_0) = \infty;\ \wp(\frac{1}{2},\Lambda_0) = e;\ \wp(\frac{1}{2}+i\frac{1}{2},\Lambda_0) = 0;\ \wp(\frac{1}{2}i,\Lambda_0) = -e;$$





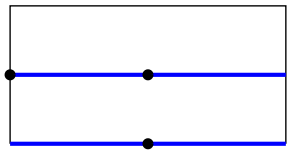


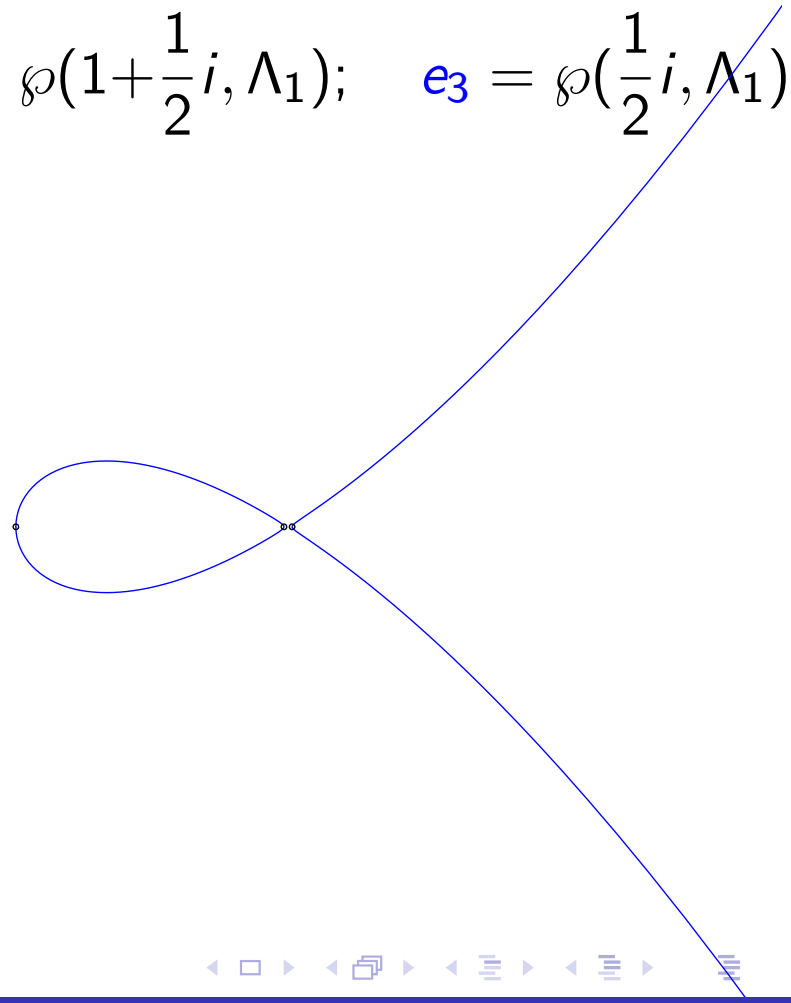
Example 1

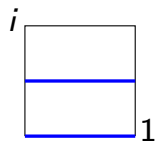
This is first iteration of agm:

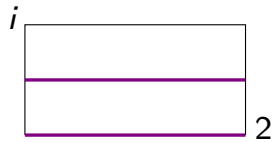
$$\Lambda_1 = 2\mathbb{Z} + i\mathbb{Z} \quad \leftrightarrow \quad y_1 = (x - e_1)(x - e_2)(x - e_3);$$

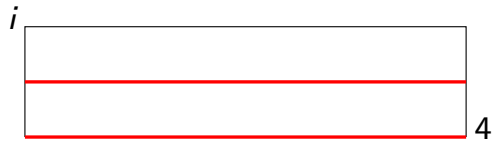
$$\infty = \wp(0, \Lambda_1);\ e_1 = \wp(1, \Lambda_1); \quad e_2 = \wp(1+\frac{1}{2}i, \Lambda_1); \quad e_3 = \wp(\frac{1}{2}i, \Lambda_1)$$

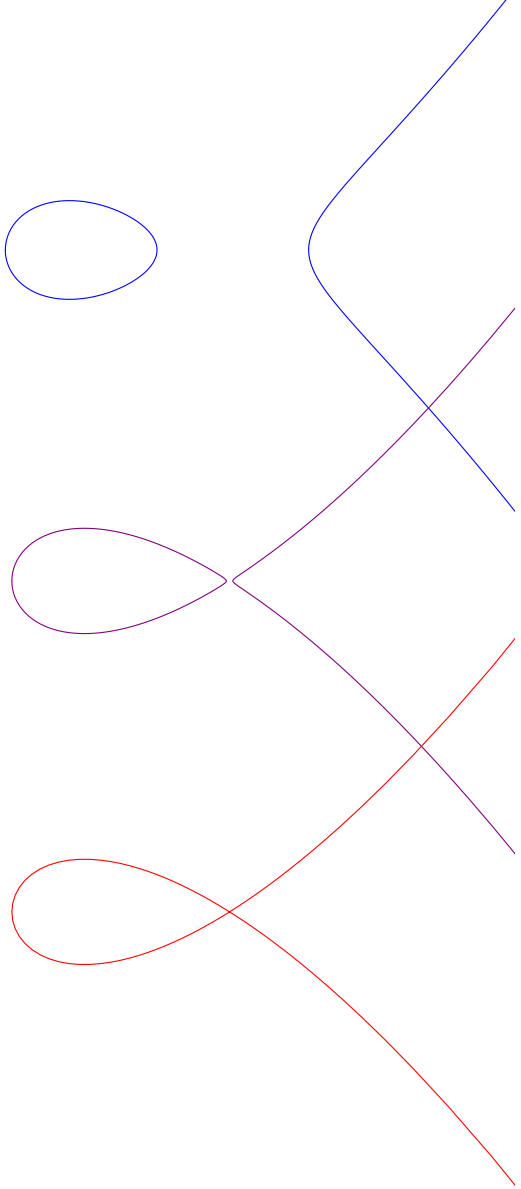





i
1


i
2


i
4



The connection between $a > b$ of agm and the roots $e_1 > e_2 > e_3$ of the elliptic curve :

$$\int_0^{\frac{\pi}{2}} \frac{1}{\sqrt{a^2(\cos\phi)^2 + b^2(\sin\phi^2)}} d\phi = \int_{e_3}^{e_2} \frac{1}{\sqrt{(x-e_1)(x-e_2)(x-e_3)}}$$

Where

$$e_1 + e_2 + e_3 = 0; \quad e_1 - e_3 = a; \quad e_2 - e_3 = b$$

Complex case

If a, b are no longer positive real numbers, but two complex numbers then we have a problem ! There is no obvious choice for

$$b_{n+1} = \sqrt{a_n b_n}.$$

We have two choices for each $n \geq 0$, so we get uncountably many possible sequences $\{a_n\}.\{b_n\}$ given a, b.

It turns out all these sequences converge, but only countably many have a non zero limit !

All possible limit values are related !

Theorem (Gauss)

There are special values μ, λ such that all possible limit values μ' are given by

$$\frac{1}{\mu'} = \frac{d}{\mu} + \frac{ic}{\lambda}$$

where d, c are relatively prime and $d \equiv 1 \mod 4$, $c \equiv 0 \mod 4$.

Let $\mathfrak{h} = \{\tau \in \mathbb{C} : \Im\tau > 0\}$ and set $q = e^{\pi i \tau}$

$$p(\tau) = 1 + 2\sum_{n=1}^{\infty} q^{n^2} = \Theta_3(\tau, 0)$$

This is a holomorphic function of τ. It is one of the Jacobi theta functions.

The set of possible values

$$\left\{ \frac{p(\gamma\tau_0)^2}{a} : \gamma \in \Gamma \right\}$$

i.e. the function $\frac{p(x)^2}{a}$ evaluated on the orbit of the point τ_0 under the group Γ.

$$\Gamma = \{\gamma \in SL_2(\mathbb{Z}); d \equiv 1(4),\ c \equiv 0(4)\}$$

Encryption

How is the agm used for encryption?

We can look at an elliptic curve E over a finite field $\mathbb{F}_q$ instead of over the complex field $\mathbb{C}$. We denote this curve $E(\mathbb{F}_q)$. It is important in this case to know the exact number of points on the curve $\#E(\mathbb{F}_q)$ (this number is $q + O(\sqrt{q})$), that is the number of $x, y \in \mathbb{F}_q$ satisfying the equation

$$y^2 = ax^3 + bx^2 + cx + d.$$

For cryptographic purposes, it is essential to find an elliptic curve E where the number $\#E(\mathbb{F}_q)$ has a large prime factor, so an estimate on its size is not good enough. The agm provides a quick way to evaluate $\#E(\mathbb{F}_q)$ (via a series of approximations to the (canonical lift of the) elliptic curve E).

Some links:

Cox, David A. *The arithmetic-geometric mean of Gauss.*. Enseign. Math. (2) 30 (1984), no. 3-4.

Ritzenthaler, Christophe *AGM for elliptic curves.*