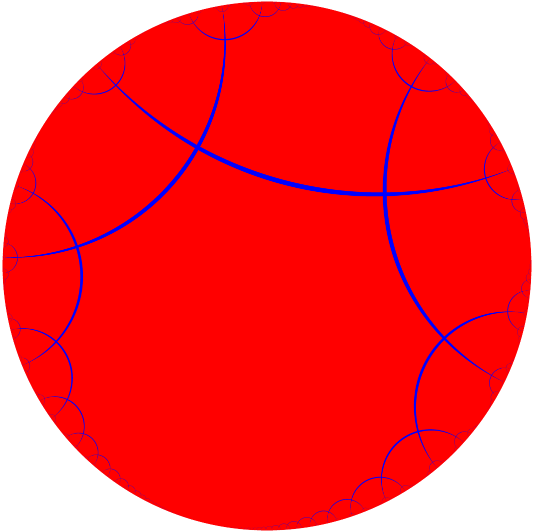
An invitation to arithmetic expression geometry

Mingli Yuan

March 5, 2024

Table of Contents

The first glimpse



The famous example of word2vec

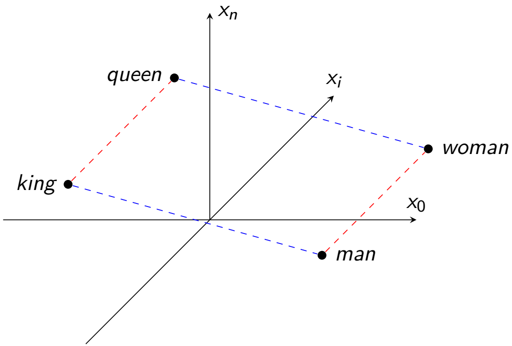


Figure: regulairty of word2vec

$$(\alpha + 1) \times 2 \neq \alpha \times 2 + 1$$

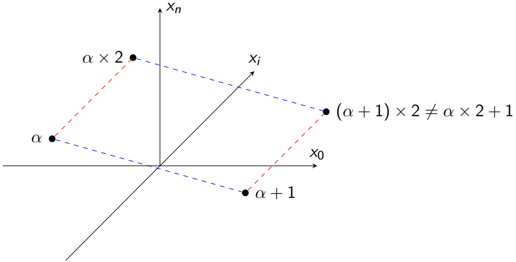
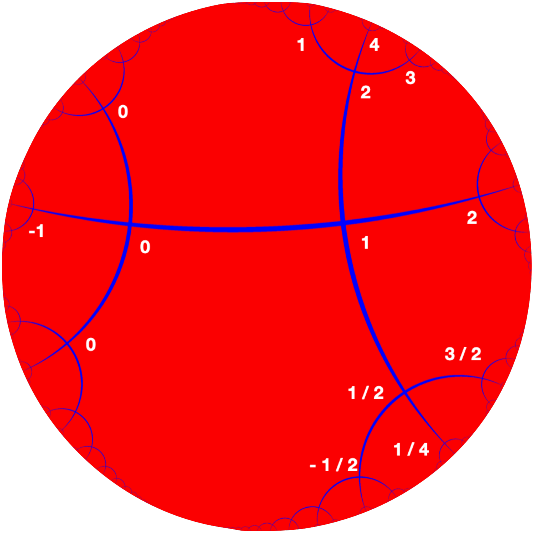


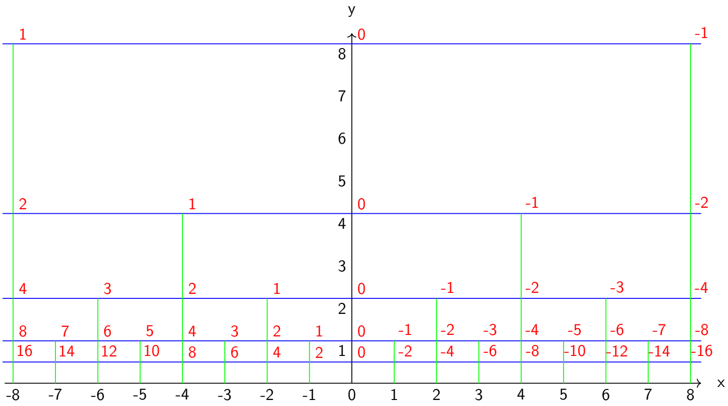
Figure: contradiction of numbers in Euclidean space

One arrangement in hyperbolic space

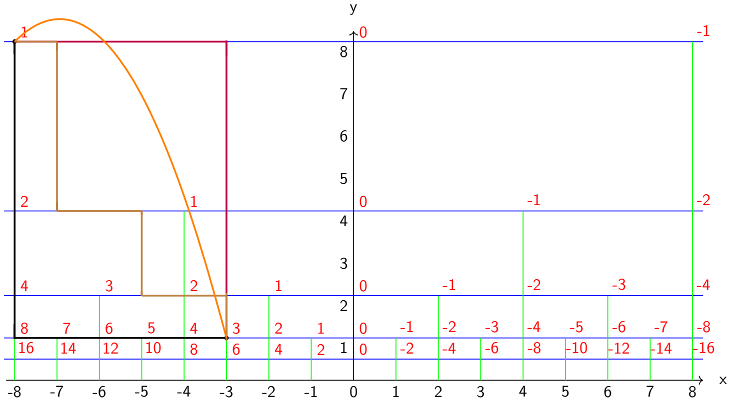


Another arrangement in hyperbolic space

$$a = -\frac{x}{y}$$



- black line $1 \times 8 - 5 = 3$



Suppose we have a base point a_0, and we step a small distance away from a_0.
Addition first

$$a_\delta = (a_0 + \mu\epsilon \cos\theta) e^{\lambda\epsilon \sin\theta}$$

Multiplication first

$$a_\delta = a_0 e^{\lambda\epsilon \sin\theta} + \mu\epsilon \cos\theta$$

Both formula can be simplified to the same result:

$$a_\delta = a_0 + \epsilon(a_0\lambda \sin\theta + \mu\cos\theta)$$

Then, we have the following equation:

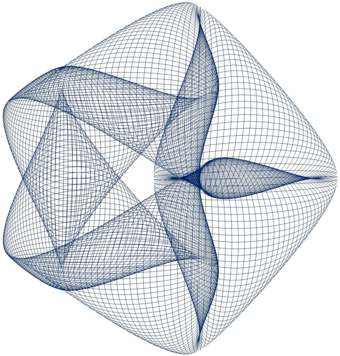
$$\frac{1}{\delta}(a_\delta - a_0) = \frac{\epsilon}{\delta}(\mu\cos\theta + a_0\lambda\sin\theta)$$

When both δ and ϵ are towards zero, we get da/dt, and hence

$$\frac{da}{dt} = u(\mu\cos\theta + a\lambda\sin\theta)$$

Or, we can change it to another form

$$\frac{da}{ds} = \mu\cos\theta + a\lambda\sin\theta \tag{1}$$



What is an arithmetic expression?

Giving an arithmetic expression, we can parse it into a syntax tree. For example, the expression

$$(((((1 \times 2) \times 2) - 1) \times (2 + 1)) - 6) \tag{2}$$

and the parsed syntax tree

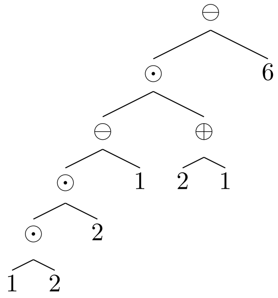


Figure: a tree representation of an arithmetic expression

Definition

An arithmetic expression a over $\mathbb{Q}$ is a structure given by the following production rules:

$$\begin{aligned} a &\longleftarrow x \\ a &\longleftarrow (a + a) \\ a &\longleftarrow (a - a) \\ a &\longleftarrow (a \times a) \\ a &\longleftarrow (a \div a) \end{aligned} \tag{3}$$

where $x \in \mathbb{Q}$, and we denote this as $a \in \mathbb{E}[\mathbb{Q}]$.

We can define evaluation $\nu(a)$ of a recursively as follows:

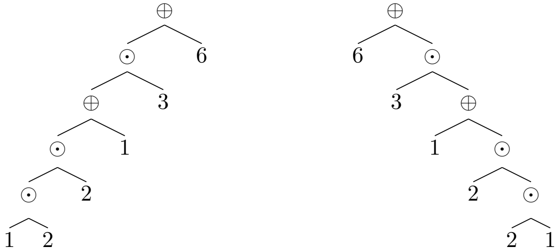
- Constant leaf: for any $x \in \mathbb{Q}$, $\nu(x) = x$.
- Compositional node by $+$: For any $(a + b)$, $\nu((a + b)) = \nu(a) + \nu(b)$.
- Compositional node by $-$: For any $(a - b)$, $\nu((a - b)) = \nu(a) - \nu(b)$.
- Compositional node by $\times$: For any $(a \times b)$, $\nu((a \times b)) = \nu(a)\nu(b)$.
- Compositional node by $\div$: For any $(a \div b)$, if $\nu(b) \neq 0$, then $\nu((a \div b)) = \nu(a)/\nu(b)$.

Generally, the evaluation order of the arithmetic expression is not unique though the result is decided.

Threadlike expressions

Right-expanded and left-expanded threadlike expressions

$$(((((1 \times 2) \times 2) + 1) \times 3) + 6) \qquad (6 + (3 \times (1 + (2 \times (2 \times 1)))))$$



The evaluation order of threadlike expressions is unique. We take right-expanded threadlike expressions as the standard form.

A careful reader may have noticed that the definition 1 is based on rational numbers $\mathbb{Q}$. Why can't we use real numbers $\mathbb{R}$ instead? The answer is that syntactically valid expressions may not be semantically valid. Dividing by zero can lead to invalid expressions, and the evaluation of the expression cannot be defined in this situation. Therefore, in real numbers, an expression may be syntactically valid but semantically not valid, and there is no algorithm that can decide whether an expression is semantically valid or not.

A two-generator group generated from

- initial operand: 0
- operator: $\oplus_\mu : x \mapsto x + \mu$
- operator: $\otimes_\lambda : x \mapsto x \cdot e^\lambda$

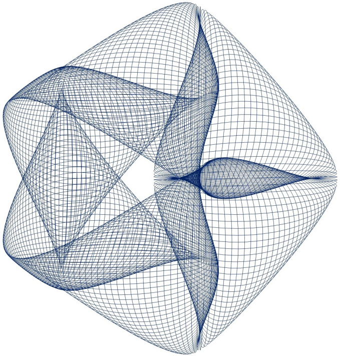
The commutator of the generators

$$x \oplus_\mu \otimes_\lambda \ominus_\mu \oslash_\lambda - x = \mu(1 - e^{-\lambda}) \tag{4}$$

$$x \otimes_\lambda \oplus_\mu \oslash_\lambda \ominus_\mu - x = -\mu(1 - e^{-\lambda}) \tag{5}$$

Arithmetic torsion reflects the non-commutativity of the group.

$$\tau = x \oplus_\mu \otimes_\lambda - x \otimes_\lambda \oplus_\mu = \mu(e^\lambda - 1) \tag{6}$$



Local polar coordinate

$$\frac{da}{ds} = \mu \cos\theta + a\lambda \sin\theta \tag{7}$$

Local Cartesian coordinate basis

$$da = \mu du + a\lambda dv \tag{8}$$

Contour-gradient coordinate

$$\frac{da}{ds} = \sqrt{\mu^2 + a^2\lambda^2} \cos\phi \tag{9}$$

where ϕ is the angle between the trace and the gradient line.

The flow equation can be written in the form of

$$da = \mu du + a\lambda dv$$

A global existence condition will limit the geometry of the space, i.e. the metric. For now we only assume the metric is well-defined and the existence condition is satisfied.

$$ds^2 = A^2 du^2 + B^2 dv^2 \tag{10}$$

The flow equation can be solved formally

$$\frac{da}{ds} = \mu \cos\theta + a\lambda \sin\theta$$

and the formal solution is

$$a = a_0 e^{\lambda s \sin\theta} + \frac{\mu}{\lambda}(e^{\lambda s \sin\theta} - 1)\cot\theta \tag{11}$$

$$a = a_0 e^{\lambda s \sin\theta} + \mu s \cos\theta + \frac{\mu}{2\lambda}\sin 2\theta\left(\frac{\lambda^2 s^2}{2!} + \frac{\lambda^3 s^3}{3!}\sin\theta + \frac{\lambda^4 s^4}{4!}\sin^2\theta + \cdots\right)$$

$$a = a_0 e^{\lambda s \sin\theta} + \mu s \cos\theta + \frac{\mu}{2\lambda}\Psi(s)\sin 2\theta$$

So $\theta = 2k\pi$ encode addition, $\theta = 2k\pi + \frac{\pi}{2}$ encode multiplication, and $\theta = 2k\pi + \pi$ encode subtraction, and $\theta = 2k\pi + \frac{3\pi}{2}$ encode division.

The flow equation can be solved formally

$$\frac{da}{ds} = \sqrt{\mu^2 + a^2\lambda^2}\cos\phi$$

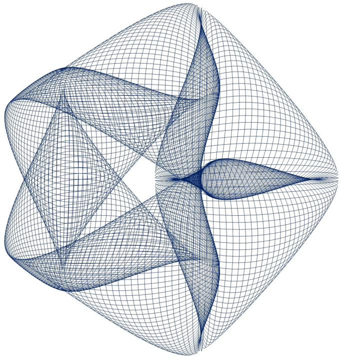
and we have

$$\tanh(\lambda s\cos\phi - c) = \frac{\lambda a}{\sqrt{\mu^2 + \lambda^2 a^2}}$$

Under the initial condition $a = 0$ when $s = 0$, we can get the following equation:

$$a = \frac{\mu}{\lambda}\sinh(\lambda s\cos\phi) \tag{12}$$

Notice the circumference of a circle is $C = 2\pi\sinh(r)$, the above equation gives us a wavefront picture.



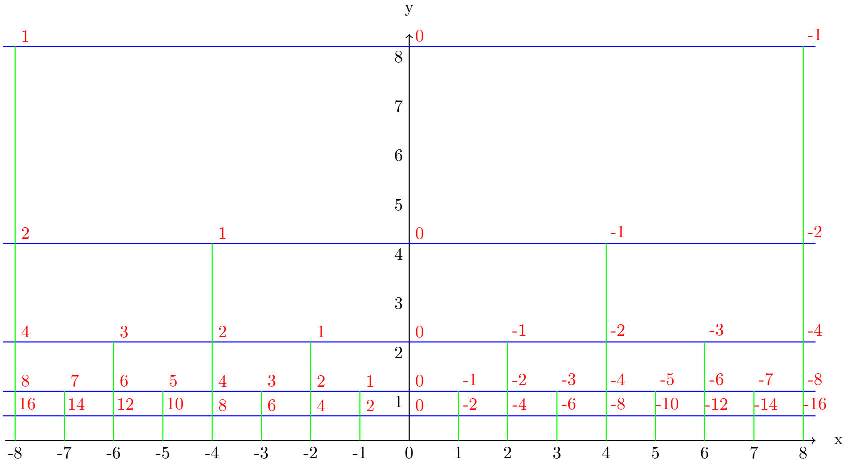
The hyperbolic space equipped with the metric

$$ds^2 = \frac{1}{y^2}\left(\frac{dx^2}{\mu^2} + \frac{dy^2}{\lambda^2}\right)$$

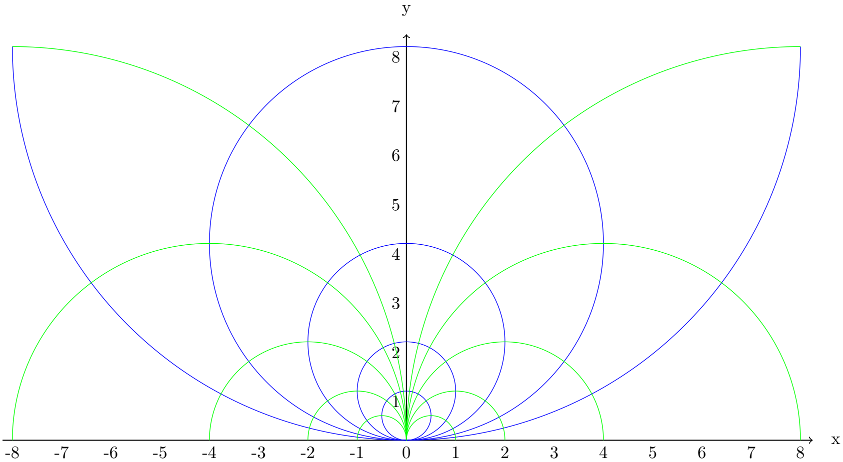
and a scalar field

$$a = -\frac{x}{y}$$

The flow equation is satisfied.


y
x
1 0 -1
2 1 0 -1 -2
4 3 2 1 0 -1 -2 -3 -4
8 7 6 5 4 3 2 1 0 -1 -2 -3 -4 -5 -6 -7 -8
16 14 12 10 8 6 4 2 0 -2 -4 -6 -8 -10 -12 -14 -16
8 7 6 5 4 3 2 1
-8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8

Another grid on $\mathfrak{E}_1$ space



In our setting, $A = \frac{1}{\mu y}$ and $B = \frac{1}{\lambda y}$:

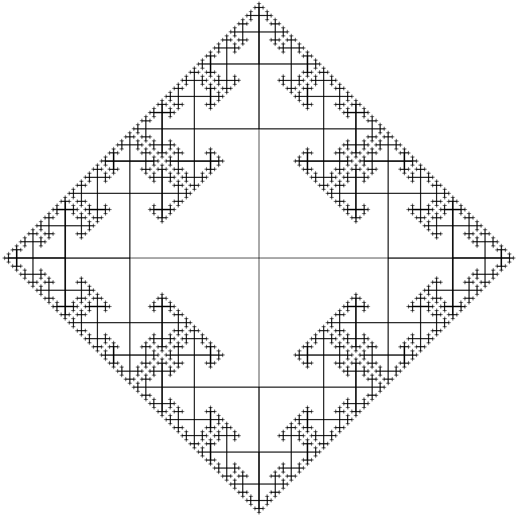
$$\Delta f = y^2 \left(\mu^2 \frac{\partial^2 f}{\partial x^2} + \lambda^2 \frac{\partial^2 f}{\partial y^2} \right)$$

And for the function $f = -\frac{x}{y}$, we have

$$\Delta f = -\frac{2\lambda^2 x}{y} = 2\lambda^2 f$$

So, we reach the conclusion that the function $f = -\frac{x}{y}$ is a eigenfunction of the Laplacian with eigenvalue $2\lambda^2$.

Further study: I - Area formula



We scale up the step size
For one step, we have

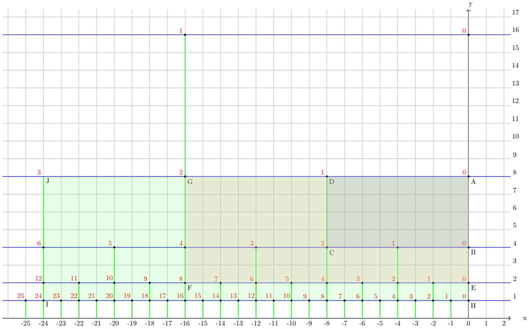
$$(x+1)\times 2-(x\times 2+1)=1 \tag{13}$$

Extending this to two steps, we encounter a different situation:

$$(x+2)\times 4-(x\times 4+2)=6 \tag{14}$$

And for three steps, the pattern continues:

$$(x+3)\times 8-(x\times 8+3)=21 \tag{15}$$



$$d\tau = (a_0 + \mu du)e^{\lambda dv} - (a_0 e^{\lambda dv} + \mu du)$$

$$d\tau = \mu\lambda dudv$$

and because

$$dS = \sqrt{A^2B^2 - F^2}dudv$$

$$\frac{d\tau}{dS} = \frac{\mu\lambda}{\sqrt{A^2B^2 - F^2}} \tag{16}$$

A broken clue?

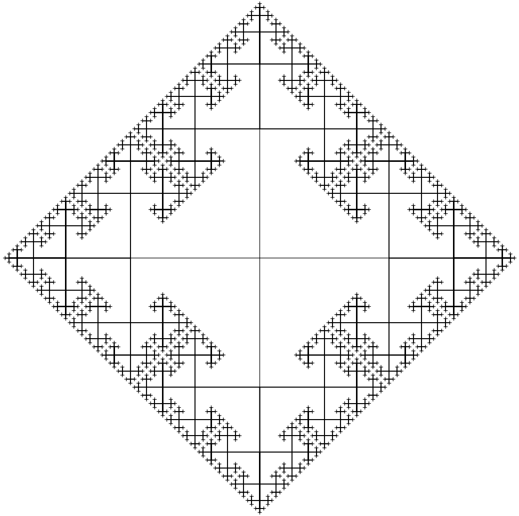
$$\begin{tikzcd}
x+\mu du \arrow[d, dashed] & & x \arrow[ll, dashed] \arrow[d, dashed] \\
(x+\mu du)e^{\lambda dv} & xe^{\lambda dv}+\mu du \arrow[l, red, "d\tau"'] & xe^{\lambda dv} \arrow[l, dashed]
\end{tikzcd}$$

$$d\tau = (x+\mu du)e^{\lambda dv} - (xe^{\lambda dv} + \mu du)$$

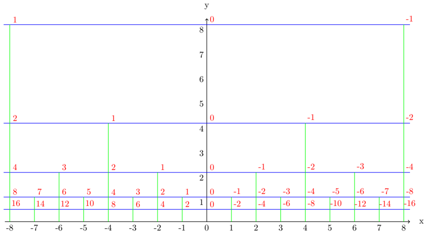
$$\begin{tikzcd}
x+\mu du \arrow[d] & & x \arrow[ll] \\
(x+\mu du)e^{\lambda dv} \arrow[r, red, "-d\tau"] & xe^{\lambda dv}+\mu du \arrow[r] & xe^{\lambda dv} \arrow[u]
\end{tikzcd}$$

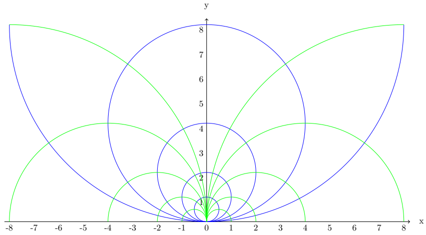
$$0 = ((x+\mu du)e^{\lambda dv} - d\tau - \mu du)/e^{\lambda dv} - x$$

Further study: II - new representation

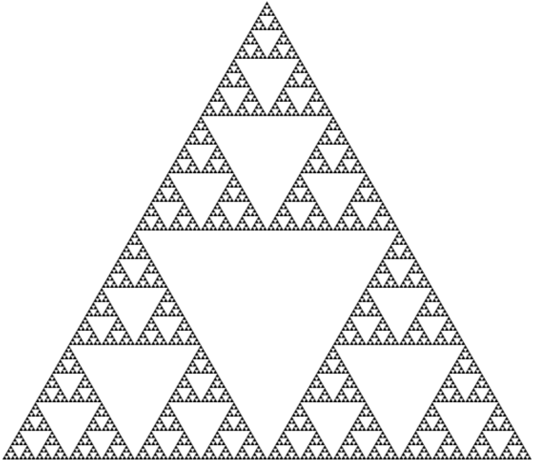


Connections between the two grids





A Rosetta stone



Topological arithmetic expression space: idea

The idea is

1. Convergence geometrically can lead to convergence of arithmetic evaluation
2. The arithmetic evaluation can be extended to the whole topological space continuously

Topological arithmetic expression space?

We denote the set of all well-defined arithmetic expressions over the field of rational numbers $\mathbb{Q}$ as $\mathbb{A}$, where ν is the evaluation function from $\mathbb{A}$ to $\mathbb{Q}$.

Definition

There is a countable dense set G on the topological space $\mathcal{A}$, and there exists an injection $\kappa : G \to \mathbb{A}$ between this dense set and the well-defined arithmetic expressions. We denote the image of this mapping as $\kappa(G) = \mathbb{K}$. If for any point $x \in \mathcal{A}$, and any two sequences of points $y_i \in G$ and $w_j \in G$, when y_i converges to x and w_j also converges to x, the sequences of points over the rational numbers $\nu(\kappa(y_i))$ and $\nu(\kappa(w_j))$ both converge to the same real number; in this case, we can naturally make an extension:

- Extend G to $\mathcal{A}$ by completeness;
- Extend ν to a mapping $\bar{\nu}$ from $\bar{\mathbb{K}}$ to the real numbers $\mathbb{R}$;

If this extended valuation function $\bar{\nu}$ is a continuous function, then we call the topological space $\mathcal{A}$ a topological arithmetic expression space. G is referred to as the grid on $\mathcal{A}$.

$\mathfrak{E}_1$ is too rigid, we need another examples. Can we find a space that is more flexible?

- Local existence theorem
- Global existence on arbitrary surface?

Local structure: totally decided by the flow equation.
Classification of the global structure

Eigenfunction of Laplacian in $\mathfrak{E}_1$ space might not be a special case.

Recall the section of Syntactic vs. Semantic, we have to consider the singularity of the space.
In complex analysis, the singularity of a function can be classified into different types, such as removable singularity, pole, and essential singularity. Similarly, we can consider the singularities in the arithmetic expression space, and their types.
This connects to the topic of divergent series in math history.

Formulate our systems as a tuple of

- $E(F), (H, a), (Path, Integ)$

Here we have

- $E(F)$: Expressions over a field F
- (H, a): A scalar field "assignment" a on a space H
- $(Path, Integ)$: all paths can be interpreted as an integral

Can we find more examples? Does complex analysis belong to this structure?

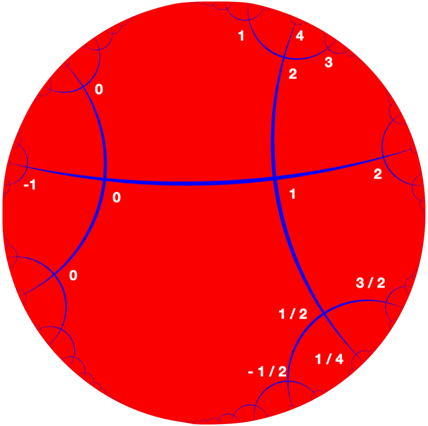
Flow perspectives

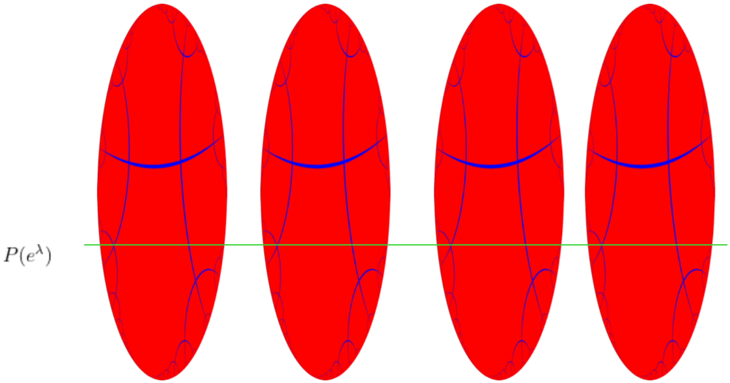
- Function as flow
- Integral as flow
- Limited process as flow

$$
\begin{tikzcd}
H \arrow[r, "I"] \arrow[d, "\nu"'] & H \arrow[d, "\nu"] \\
R \arrow[r, "k"'] & R
\end{tikzcd}
$$

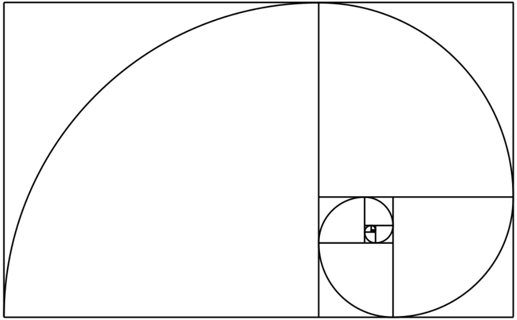
Riemann sum is purely additional. Can we extend it by mixing addition and multiplication?

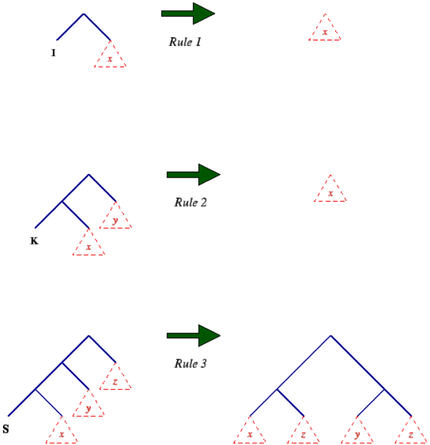
Infinitely small values and large values are appeared at the boundary. Limitation process can be treated as flow to the boundary.




$P(e^\lambda)$

Hyper-operation
Flow treatment of binary operation
Higher-dimensional space





S(K(SI))Kαβ →

K(SI)α(Kα)β →

SI(Kα)β →

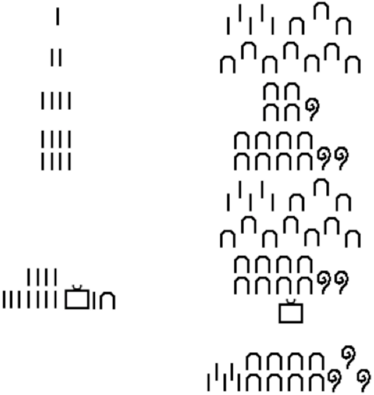
Iβ(Kαβ) →

Iβα →

βα

A space of SKI combinators

Any arithmetic expression can be represented by SKI combinators via Church numerals.
So any arithmetic expression space can be encoded by SKI combinators.
Program space!



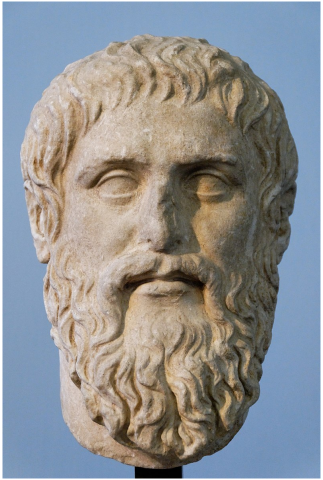
a example calculation

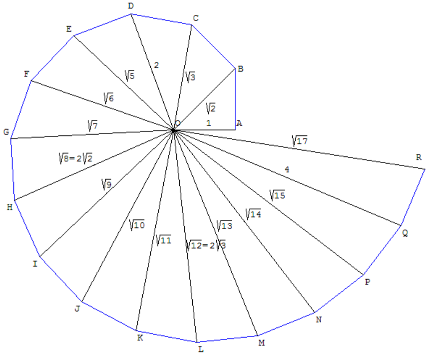
1*	35
2*	70
4	140
8*	280
1+2+8=11	35+70+280=385

Ahmes Papyrus

http://www.math.buffalo.edu/mad/Ancient-Africa/mad_ancient_egypt_arith.html

A story from Plato's Theaetetus: Theodorus of Cyrene, a young mathematician, was able to prove that $\sqrt{3}$, $\sqrt{5}$... are irrational, but not $\sqrt{17}$.





The axioms by Victor Pambuccian and Celia Schacht are also fit into expressions, and then some terms of the system can be embedded into the expression space. Can we migrate the problem of irrationality of $\sqrt{17}$ from proof theory into a problem in group theory?

A 1. $(x + y) + z = x + (y + z)$

A 2. $x + y = y + x$

A 3. $(x \cdot y) \cdot z = x \cdot (y \cdot z)$

A 4. $x \cdot y = y \cdot x$

A 5. $x \cdot (y + z) = x \cdot y + x \cdot z$

A 6. $x + 0 = x \wedge x \cdot 0 = 0$

A 7. $x \cdot 1 = x$

A 8. $(x < y \wedge y < z) \to x < z$

A 9. $\neg x < x$

A 10. $x < y \vee x = y \vee y < x$

A 11. $x < y \to x + z < y + z$

A 12. $(0 < z \wedge x < y) \to x \cdot z < y \cdot z$

A 13. $x < y \to x + (y - x) = y$

A 14. $0 < 1 \wedge (x > 0 \to (x > 1 \vee x = 1))$

A 15. $x > 0 \vee x = 0$

A 16. $m = \kappa(m, n) \cdot \mu(m, n) \wedge n = \kappa(m, n) \cdot \mu(n, m) \wedge (\mu(m, n) = 2\left[\frac{\mu(m,n)}{2}\right] + 1 \vee \mu(n, m) = 2\left[\frac{\mu(n,m)}{2}\right] + 1).$

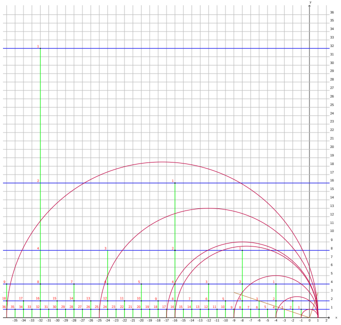
A 17. $x = \left[\frac{2x}{2}\right].$

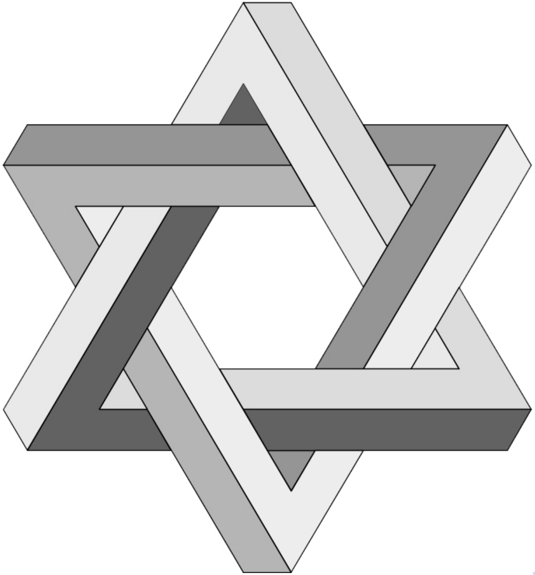
A 18. $x = 2\left[\frac{x}{2}\right] \vee x = 2\left[\frac{x}{2}\right] + 1$

A 19. $\pi_2(n) \wedge a \cdot b = n \wedge a > 1 \to a = 2\left[\frac{a}{2}\right]$

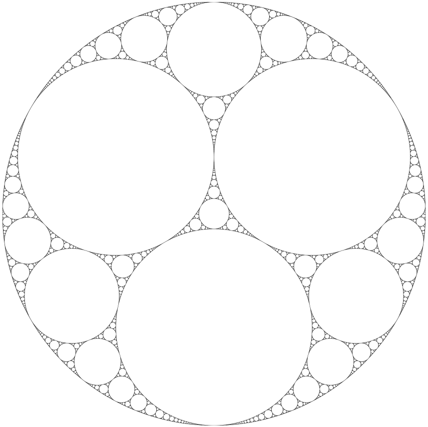
A 20. $0 < n \to n = \tau(n) \cdot \omega(n) \wedge \pi_2(\tau(n)) \wedge \omega(n) = 2\left[\frac{\omega(n)}{2}\right] + 1$

A 21. $n < m \wedge \pi_2(m) \wedge \pi_2(n) \to \tau(m - n) = n.$

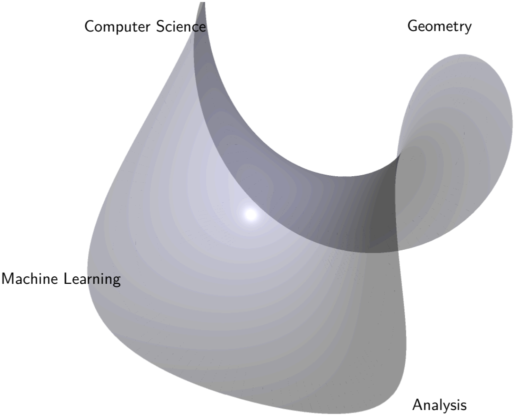




Math and even all human knowledge is also a geometrical object, just the same as the universe.



A minimal surface of knowledge, every concept is a point, every relation is a line.



Thank you!