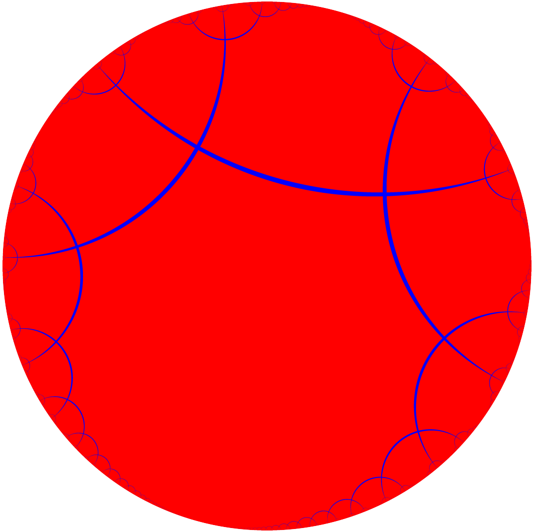
Arithmetic expression geometry with an application on learnable non-linearity

Mingli Yuan

May 24, 2023



The famous example of word2vec

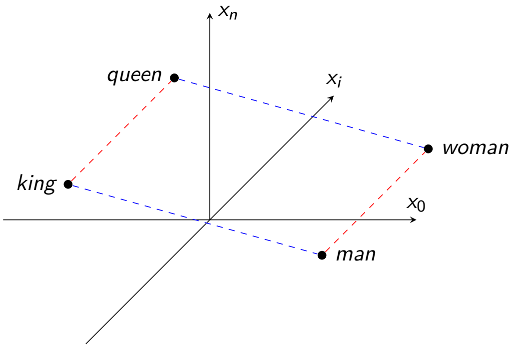


Figure: regulairty of word2vec

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

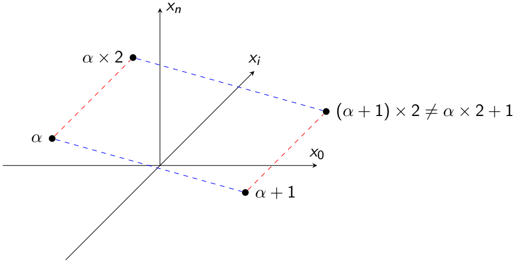
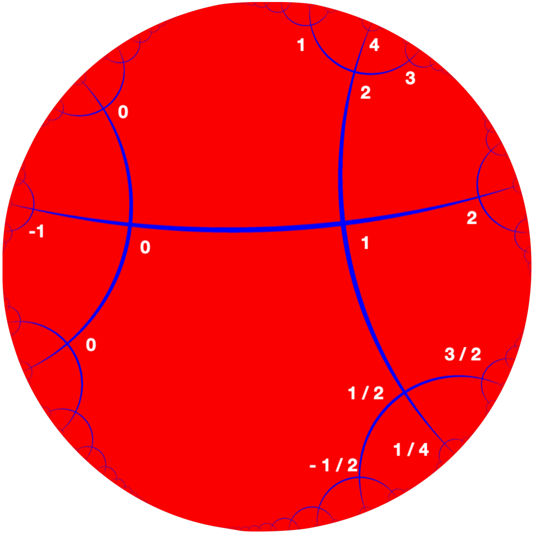
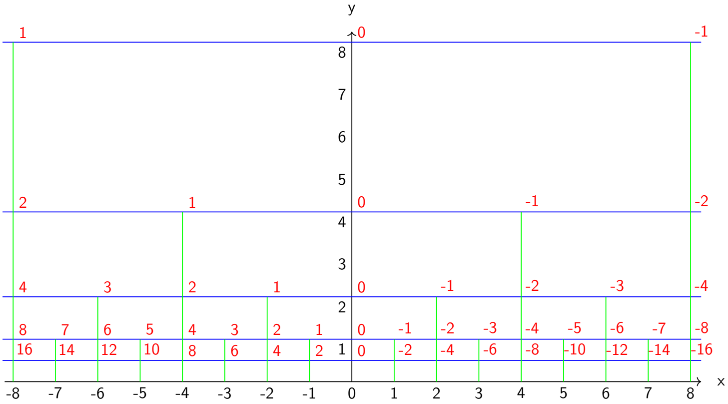


Figure: contradiction of numbers in Euclidean space

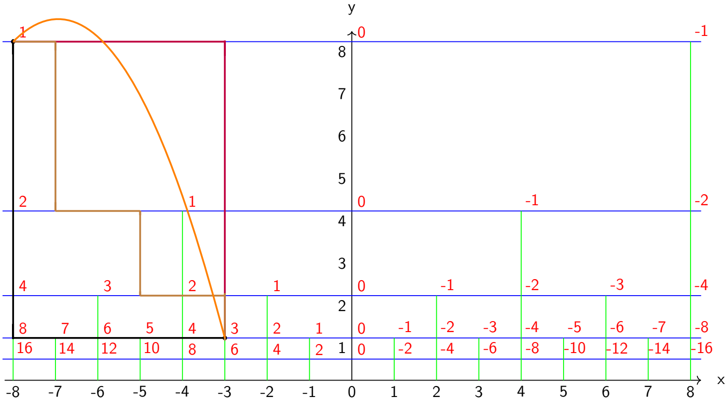
One arrangement in hyperbolic space



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


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

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



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

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

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:

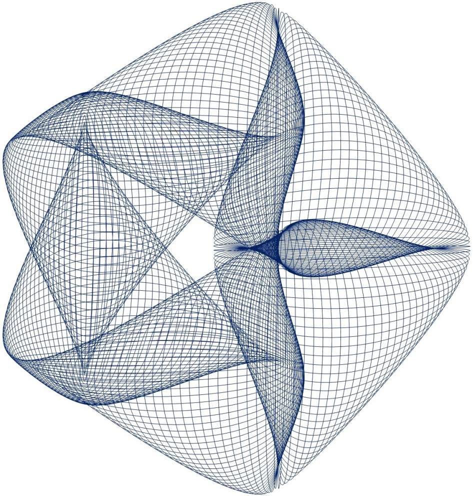
$$\frac{1}{\delta}(a_\delta - a_0) = \frac{\epsilon}{\delta}(\mu\cos\theta + x_0\lambda\sin\theta)$$

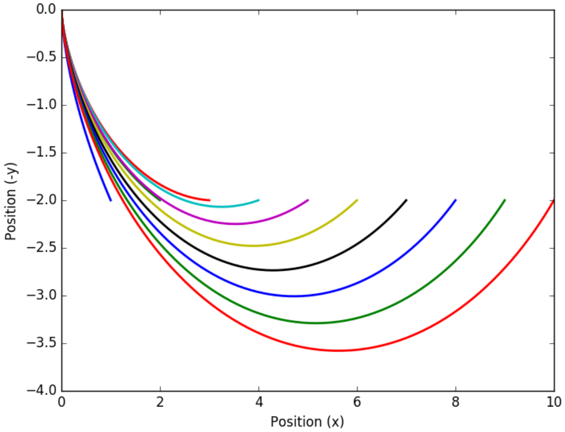
When both δ and ϵ are towards zero, we get da/dt, and hence

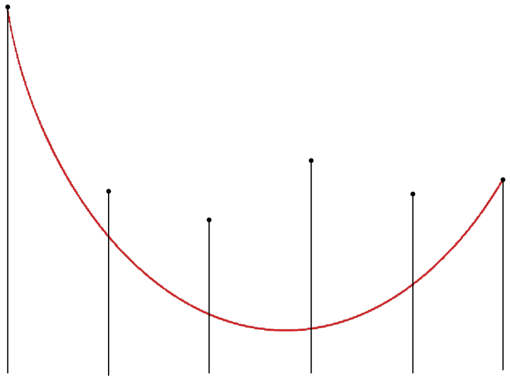
$$\frac{da}{dt} = u(\mu\cos\theta + a\lambda\sin\theta)$$

Or, we can change it to another form

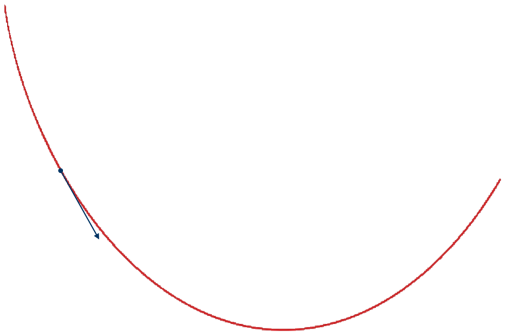
$$\frac{da}{ds} = \mu\cos\theta + a\lambda\sin\theta \tag{1}$$



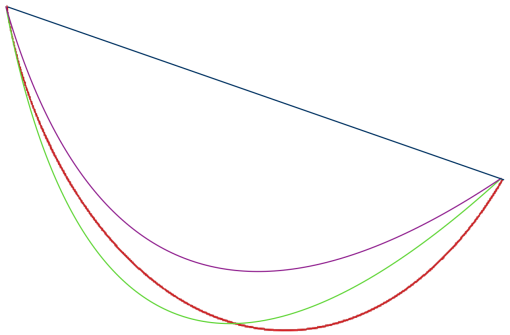

0.0
−0.5
−1.0
−1.5
−2.0
−2.5
−3.0
−3.5
−4.0
Position (-y)
0
2
4
6
8
10
Position (x)

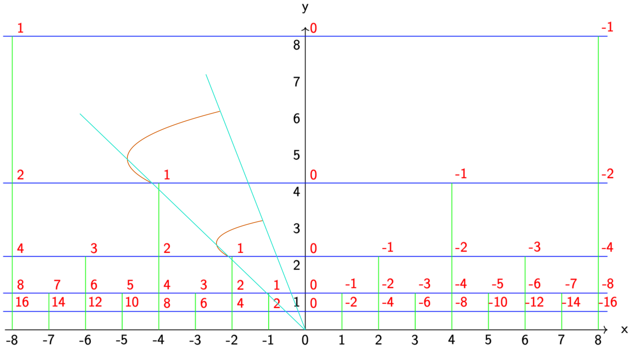


An improvement

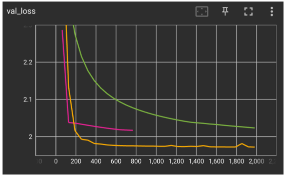


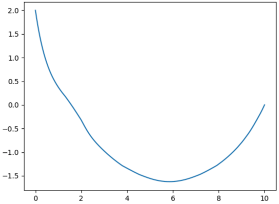
Trap




y
x
-8
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8

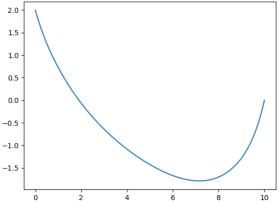
Experiment results

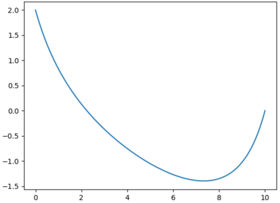




Learnable non-linearity
+ tensor linearity

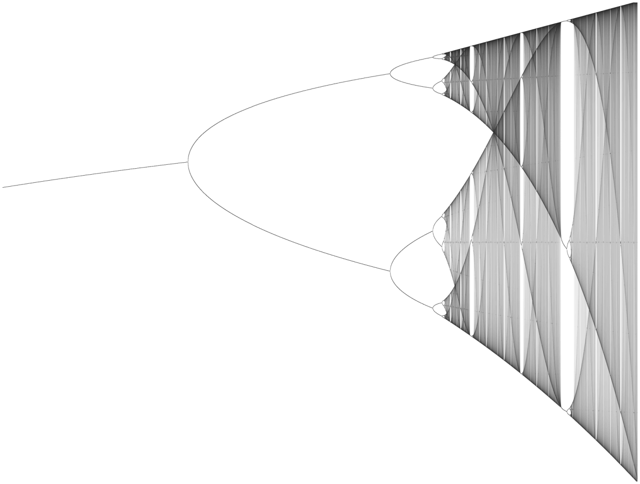
LSTM





Learnable non-linearity
+ matrix linearity

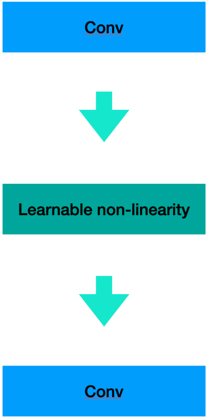
A learnable non-linearity



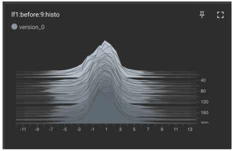
Experiment results

	logistic map p = 3.5	logistic map p = 3.8	logistic map p = 3.9	spline	learnable non-linearity
task success	✗	✓	✓	✗	✓
parameter learnable	✗	✗	✗	✗	✓

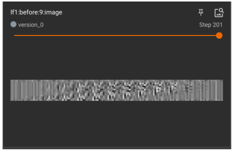
Hebbian?



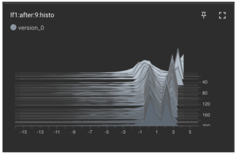
Conv
Learnable non-linearity
Conv



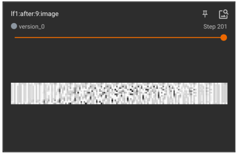
lf1:before:9:histo
version_0



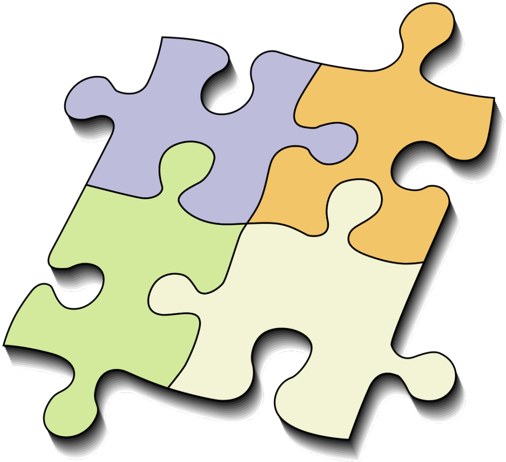
lf1:before:9:image
version_0
Step 201



lf1:after:9:histo
version_0



lf1:after:9:image
version_0
Step 201



On the hyperbolic plane

$$ds^2 = \frac{dx^2 + dy^2}{y^2}$$

the Laplacian is

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

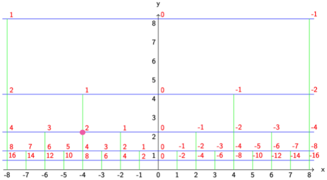
Given

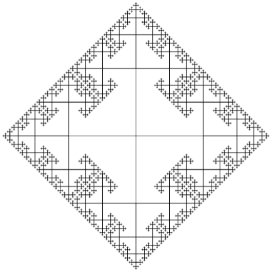
$$A = -\frac{x}{y} \tag{2}$$

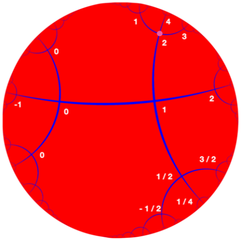
We have

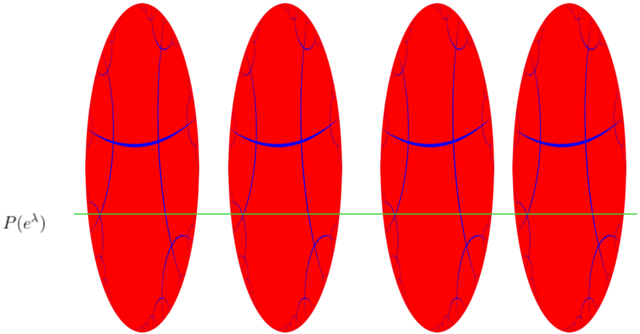
$$\Delta A = -y^2\left(\frac{\partial^2}{\partial x^2}A + \frac{\partial^2}{\partial y^2}A\right) = y^2\left(\frac{1}{\partial y}\left(\frac{1}{\partial y}\frac{x}{y}\right)\right) = 2A$$

A jigsaw by Riemann mapping








$P(e^{\lambda})$