

# A New Dragon in the Den

*Fast Conversion from Floating Point  
Numbers*

Cassio Neri

2024

# Hercules slaying the Hydra

---



Hercules slaying the Hydra

Louis Chéron

Oil painting, ca. 1690-1725

Victoria & Albert Museum, London, UK

# Hercules slaying the Hydra

---



Hercules slaying the Hydra

Louis Chéron

Oil painting, ca. 1690-1725

Victoria & Albert Museum, London, UK

*Bicho de sete cabeças.*

Seven-headed beast.

# Hercules slaying the Hydra



Hercules slaying the Hydra  
Louis Chéron  
Oil painting, ca. 1690-1725  
Victoria & Albert Museum, London, UK

*Bicho de sete cabeças.*

Seven-headed beast.



Album The Number of the Beast  
Iron Maiden, 1982.



Medieval French Apocalypse Tapestry, Produced between 1377 and 1382.  
Château d'Angers, France  
Credits: Jean-Pierre Dalbéra, CC BY 2.0, via Wikimedia Commons.

Scream for me  
Aspen !!!



Bruce Dickinson in Brno, Czech Republic, June 8, 2014  
Credits: Vaclav Salek/CTK Photo/Alamy Live News

# Hercules slaying the Hydra

---



Hercules slaying the Hydra

Louis Chéron

Oil painting, ca. 1690-1725

Victoria & Albert Museum, London, UK

*Bicho de sete cabeças.*

Seven-headed beast.

# Hercules slaying the Hydra

---



Hercules slaying the Hydra

Louis Chéron

Oil painting, ca. 1690-1725

Victoria & Albert Museum, London, UK

*Bicho de sete cabeças.*

Seven-headed beast.

*Não é um bicho de sete cabeças.*

It's not a seven-headed beast.

# Hercules slaying the Hydra

---



Hercules slaying the Hydra

Louis Chéron

Oil painting, ca. 1690-1725

Victoria & Albert Museum, London, UK

*Bicho de sete cabeças.*

Seven-headed beast.

*Não é um bicho de sete cabeças.*

It's not a seven-headed beast.

It's not rocket science.

# A New Dragon in the Den

Fast conversion from floating-point numbers

**Cassio Neri**  
(Independent Researcher)

C++Now 2024 - Aspen

# π

---



<https://godbolt.org/z/a8qETqzbj>

```
① sprintf(s1.data(), "%g", pi); // 3.14159
② snprintf(s2.data(), s2.size(), "%g", pi); // 3.14159
    string s3 = (stringstream{} << pi).str(); // 3.14159
    string s4 = to_string(pi); // 3.141593
    to_chars(s5.data(), s5.data() + s5.size(), pi); // 3.141592653589793
    string s6 = format("{}", pi); // 3.141592653589793
```

# π



```
① sprintf(s1.data(), "%g", pi);           // 3.14159
② snprintf(s2.data(), s2.size(), "%g", pi); // 3.14159
    string s3 = (stringstream{} << pi).str();   // 3.14159
    string s4 = to_string(pi);                  // 3.141593
    to_chars(s5.data(), s5.data() + s5.size(), pi); // 3.141592653589793
    string s6 = format("{}", pi);                // 3.141592653589793
```

<https://godbolt.org/z/a8qETqzbj>



```
s1 = f'{pi}'           // 3.141592653589793
s2 = str(pi)           // 3.141592653589793
s3 = pi.__str__()      // 3.141592653589793
s4 = "{}".format(pi)   // 3.141592653589793
s5 = "%s" % pi         // 3.141592653589793
```

<https://godbolt.org/z/5aseWr7fz>



<https://onecompiler.com/javascript/428wnkarf>

```
const s1 = "" + pi;          // 3.141592653589793
const s2 = pi.toString();    // 3.141592653589793
const s3 = String(pi);       // 3.141592653589793
const s4 = `${pi}`;          // 3.141592653589793
```



```
String s1 = "" + pi;          // 3.141592653589793
String s2 = Double.toString(pi); // 3.141592653589793
String s3 = String.valueOf(pi); // 3.141592653589793
String s4 = String.format("%g", pi); // 3.14159
```

<https://godbolt.org/z/8M483Kvba>



<https://godbolt.org/z/Y46WY4ooY>

```
let s1 = pi.to_string();      // 3.141592653589793
let s2 = format!("{}", pi);   // 3.141592653589793
```

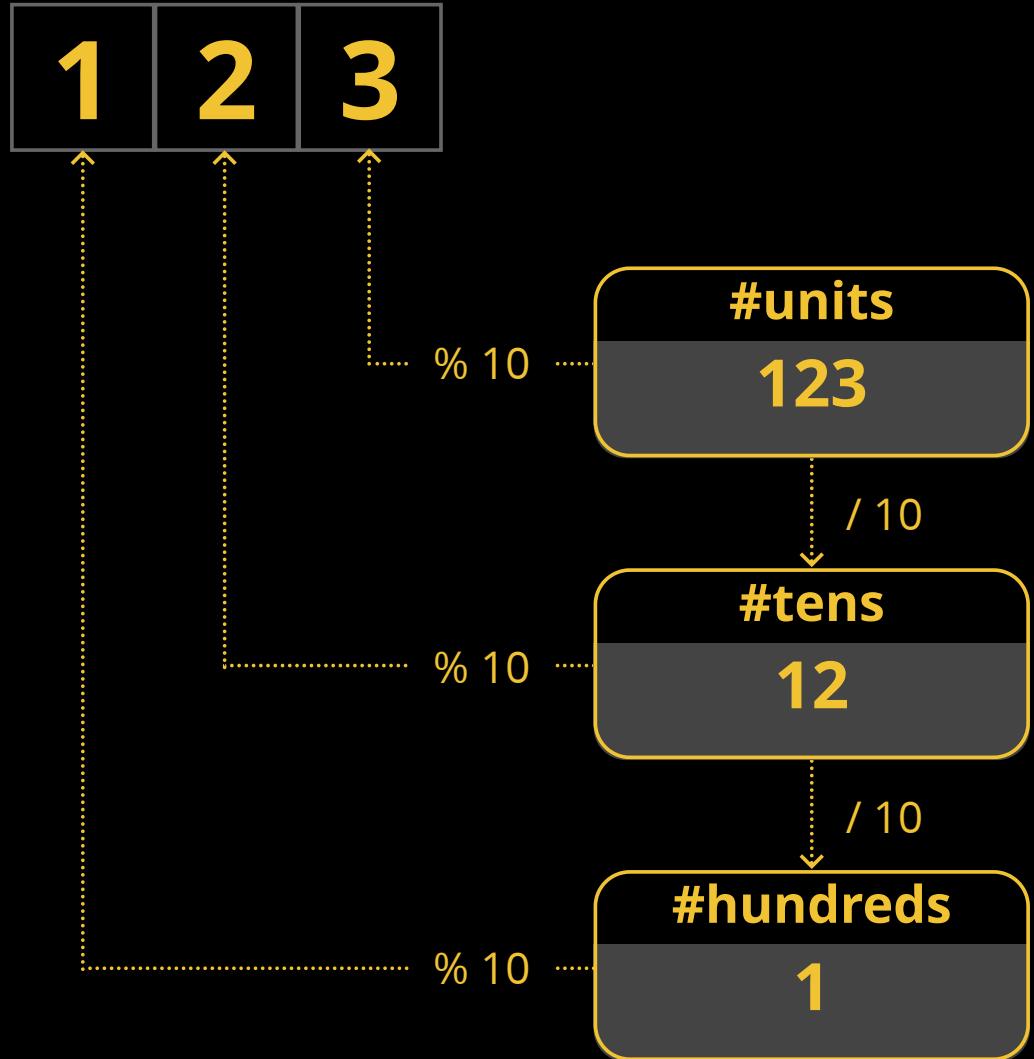


<https://godbolt.org/z/czhrT1hhh>

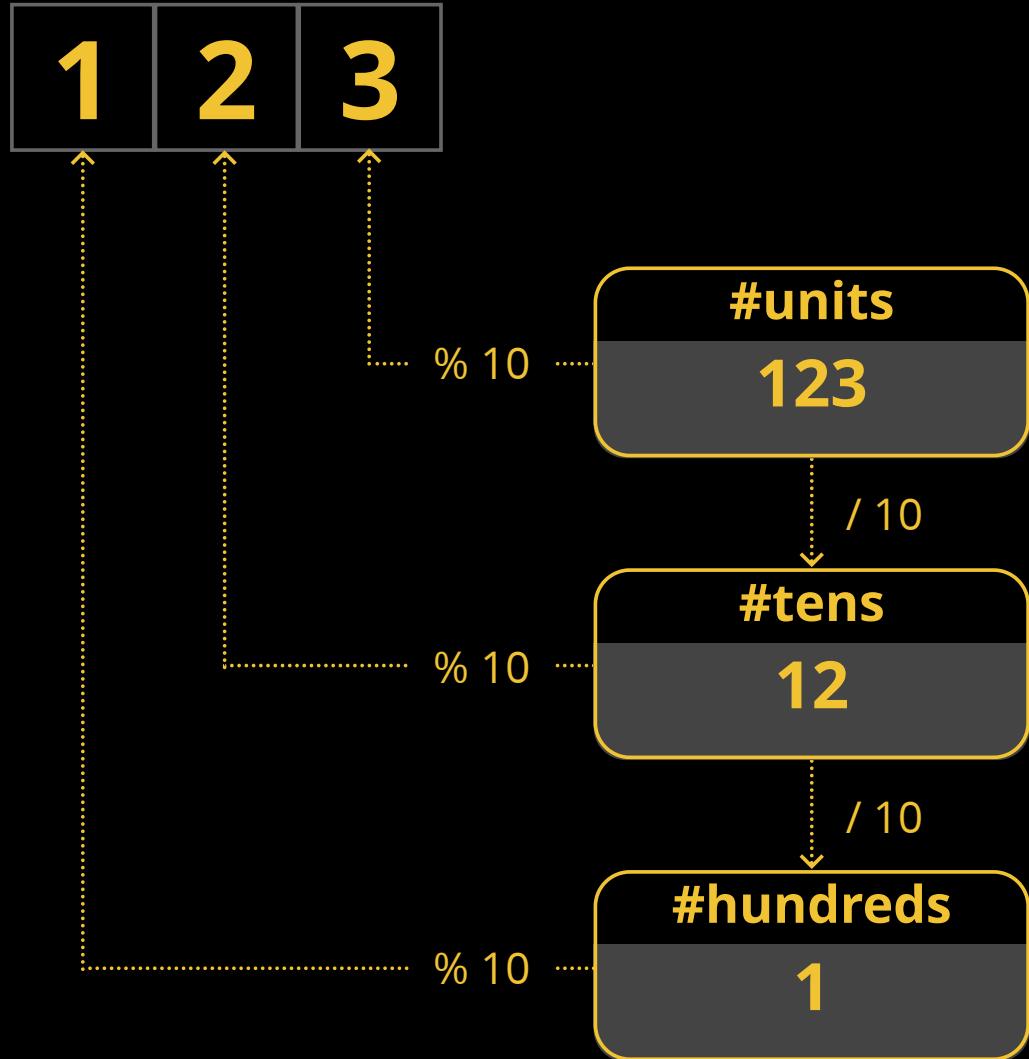
```
string s1 = pi.ToString();      // 3.141592653589793
string s2 = string.Format("{0}", pi); // 3.141592653589793
string s3 = $"{pi}";             // 3.141592653589793
```

# Integer to string

---



# Integer to string



<https://godbolt.org/z/1G9hbKMnP>

```
string convert(unsigned n) {  
  
    size_t size = number_of_digits(n);  
    string str(size, '\0');  
    char* p = &str.back();  
  
    do {  
  
        *p = n % 10 + '0';  
        n /= 10;  
  
        --p;  
    } while(n);  
  
    return str;  
}
```

## Minimum strictly positive float

---

$$2^{-149} = 0.000,000,000,000,000,000,000,000,000,  
000,000,000,001,401,298,464,324,817,  
070,923,729,583,289,916,131,280,261,941,  
876,515,771,757,068,283,889,791,082,685,  
860,601,486,638,188,362,121,582,031,25$$

(149 digits after the dot)



## Minimum strictly positive float

---

$$2^{-149} = 0.000,000,000,000,000,000,000,000,000,  
000,000,000,001,401,298,464,324,817,  
070,923,729,583,289,916,131,280,261,941,  
876,515,771,757,068,283,889,791,082,685,  
860,601,486,638,188,362,121,582,031,25$$

(149 digits after the dot)



$$\approx 10^{-45}$$



# Mechanical counter

---



# Mechanical counter

---



# Mechanical counter

---



## Decimal fixed-point representation

---

1234

## Decimal fixed-point representation

---

1234

# Decimal floating-point representation

---

exponent  
10<sup>1</sup> x 2.34 mantissa

# Decimal floating-point representation

# exponent

$10^1 \times 2.34$  mantissa

## mantissa

0,000,000,023.40

# Decimal floating-point representation

---

exponent  
 $10^1 \times 2.34$  mantissa

0 , 0   0 , 0   0 , 0   2   3 . 4   0

# Decimal floating-point representation

---

exponent  $\neq 0$

$\Rightarrow$

mantissa's first digit  $\neq 0$   
(scientific notation)

exponent  
 $10^1 \times 2.34$  mantissa

0 , 0   0 , 0   0 , 0   2   3 . 4   0

# Decimal floating-point representation

# exponent

## mantissa

$10^9 \times 9.99 \times 10^{-3}$  mantissa

9,990,000.000,00

# Decimal floating-point representation

---

biased exponent

$10^{9-3} \times 9.99$

mantissa

9 , 9   9   0 , 0   0   0 . 0   0   0 , 0   0

biased exponent  $\neq 0$

$\Rightarrow$

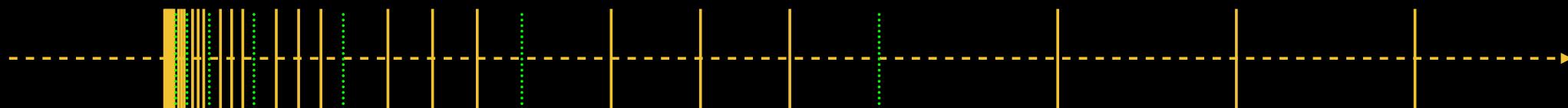
mantissa's first digit  $\neq 0$   
(scientific notation)

# Binary floating-point representation

---

Exponent and mantissa are in binary.

biased exponent  
2<sup>000-11</sup> × 0.00



# Binary floating-point representation

---

Exponent and mantissa are in binary.

biased exponent  $\neq 0$



mantissa's first bit = 1  
(normal)

biased exponent  
**000-11** **x.000**



# Binary floating-point representation

---

Exponent and mantissa are in binary.

biased exponent  $\neq 0$



mantissa's first bit = 1  
(normal)

2 **000-11** x 0.00

— Centred value

..... Uncentred value



# Binary floating-point representation

Exponent and mantissa are in binary.

# biased exponent $\neq 0$

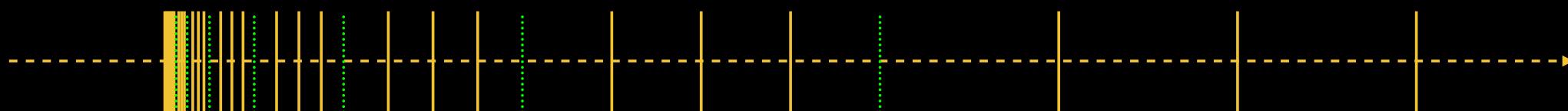
1

mantissa's first bit = 1  
(normal)

biased exponent  
2 000-11 x 1.00 mantissa

— Centred value

..... Uncentred value



# IEEE-754 representation

---

- *Binary32* (`float` and `std::float32_t`)
- *Binary64* (`double` and `std::float64_t`)
- *Binary128* (`__float128` and `std::float128_t`)
- Others

# Binary32

# Binary32

# Binary32

# Binary32

# Binary32

**0** sign: + (0) or - (1)

# Binary32

+

# Binary32

**1 0 0 0 0 0 1 0**    biased exponent  $\neq 0 \Rightarrow$  normal

$$\begin{array}{r} - \ 0 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \\ \hline 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 1 \end{array} \quad \text{bias: } 01111111 \text{ (127, normal) or } 01111110 \text{ (126, subnormal)}$$

+

# Binary32

biased exponent  $\neq 0 \Rightarrow$  normal

$$+ 2^{\textcolor{blue}{00000011}} \times 1.$$

# Binary32

$$+2^{00000011} \times 1.010110000000000000000000000000$$

# Binary32

$$+2^3 x \quad 1.3 \ 4 \ 3,7 \ 5$$

# Binary32

---

±	biased exponent	mantissa
1	8	23
0 1 0 0 0 0 1 0 0 1 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		

+2<sup>-00010100</sup> x 10101100000000000000000000000000

$\oplus 2^{-20} \times (1 \ 1,2 \ 7 \ 2,1 \ 9 \ 2)$

# Quiz

---

What's the output?



<https://godbolt.org/z/EzWTGEoP8>

```
double x = 1.0 / 3.0;  
std::cout << std::format("{}", x);
```

Options:

- (a) 0.3
- (b) 0.33333333333333333
- (c) 0.33333333333333334
- (d) One third
- (e) *None of the above*

# Quiz

---

What's the output?



<https://godbolt.org/z/EzWTGEoP8>

```
double x = 1.0 / 3.0;  
std::cout << std::format("{}", x);
```

Options:

(a) 0.3

(b) 0.3333333333333333**3**

(c) 0.3333333333333333**4**

(d) One third

(e) *None of the above*

# Quiz

---

What's the output?



<https://godbolt.org/z/EzWTGEoP8>

```
double x = 1.0 / 3.0;  
std::cout << std::format("{}", x);
```



<https://godbolt.org/z/4qP56b6Ks>

```
float x = 1.f / 3.f;  
std::cout << std::format("{}", x);
```

Options:

(a) 0.3

(b) 0.33333333333333333**3**

(c) 0.3333333333333333**4**

(d) One third

(e) *None of the above*

(a) 0.3

(b) 0.333333**3**

(c) 0.333333**4**

(d) One third

(e) *None of the above*

# Quiz

---

What's the output?



<https://godbolt.org/z/EzWTGEoP8>

```
double x = 1.0 / 3.0;  
std::cout << std::format("{}", x);
```



<https://godbolt.org/z/4qP56b6Ks>

```
float x = 1.f / 3.f;  
std::cout << std::format("{}", x);
```

Options:

(a) 0.3

(b) 0.33333333333333333**3**

(c) 0.33333333333333333**4**

(d) One third

(e) *None of the above*

(a) 0.3

(b) 0.333333**3**

(c) 0.333333**4**

(d) One third

(e) *None of the above*

# Quiz

---

What's the output?



<https://godbolt.org/z/EzWTGEoP8>

```
double x = 1.0 / 3.0;  
std::cout << std::format("{}", x);
```



<https://godbolt.org/z/4qP56b6Ks>

```
float x = 1.f / 3.f;  
std::cout << std::format("{}", x);
```

Options:

(a) 0.3

(b) 0.33333333333333333

(c) 0.33333333333333334

(d) One third

(e) *None of the above*

(a) 0.3

(b) 0.33333333333333334

(c) 0.33333333333333334

(d) One third

(e) *None of the above*



<https://godbolt.org/z/17jjd6Kv1>



<https://godbolt.org/z/43EnYG6Pj>



<https://godbolt.org/z/Yax6xdYf3>

$\frac{1}{3}$

---

↑

— 0.333,333,333,333,333,333,333,333,3



$\frac{1}{3}$

---

$$0.\overline{3} = 2^{-25} \times (11,184,810 + \frac{2}{3})$$



$\frac{1}{3}$

---

$$0.333,333,343,267,440,795,898,437,5 = 2^{-25} \times 11,184,811$$

$$0.333,333,333,333,333,333,333,333,3 = 2^{-25} \times (11,184,810 + \frac{2}{3})$$


$$0.333,333,313,465,118,408,203,125,0 = 2^{-25} \times 11,184,810$$

$\frac{1}{3}$

---

$$0.333,333,343,267,440,795,898,437,5 = 2^{-25} \times 11,184,811$$

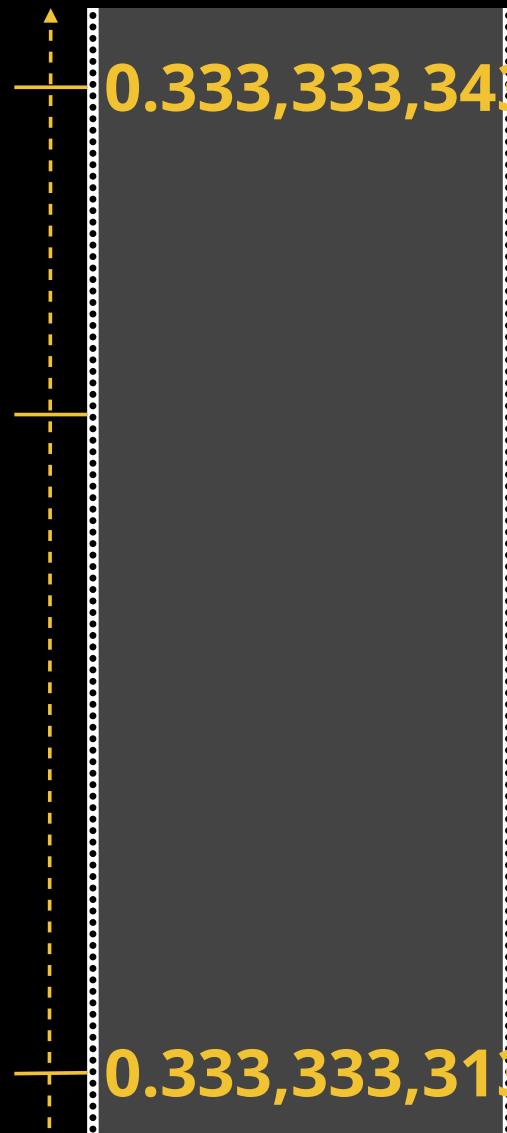


$$0.333,333,333,333,333,333,333,333,3 = 2^{-25} \times (11,184,810 + \frac{2}{3})$$

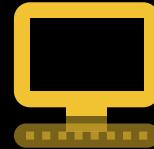


$$0.333,333,313,465,118,408,203,125,0 = 2^{-25} \times 11,184,810$$

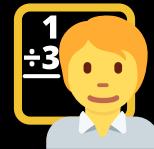
$\frac{1}{3}$



$$0.333,333,343,267,440,795,898,437,5 = 2^{-25} \times 11,184,811$$

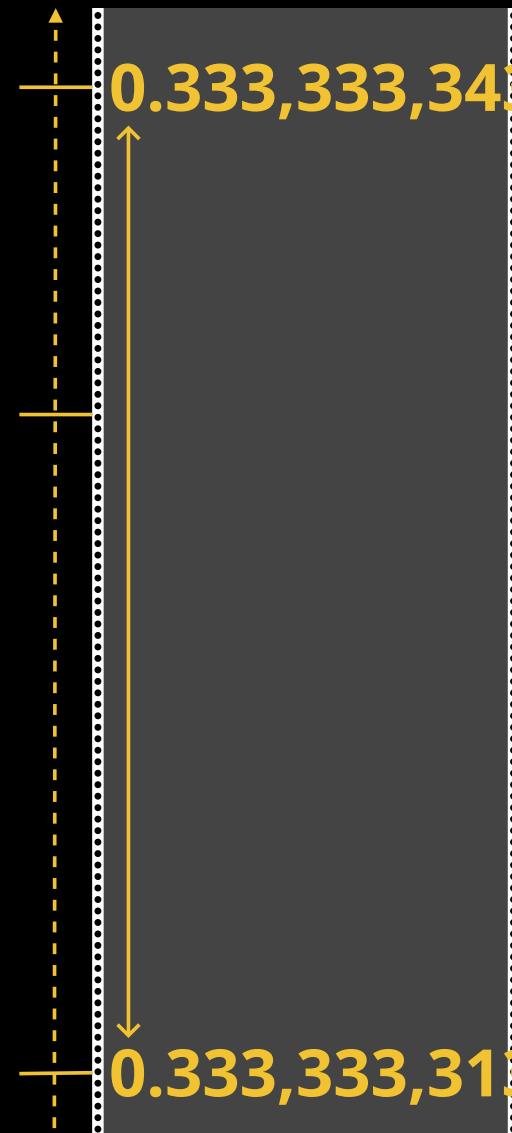


$$3,333,333,333,333,333,3 = 2^{-25} \times (11,184,810 + \frac{2}{3})$$

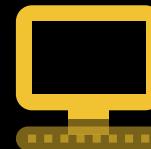


$$0.333,333,313,465,118,408,203,125,0 = 2^{-25} \times 11,184,810$$

$\frac{1}{3}$



$$0.333,333,343,267,440,795,898,437,5 = 2^{-25} \times 11,184,811$$

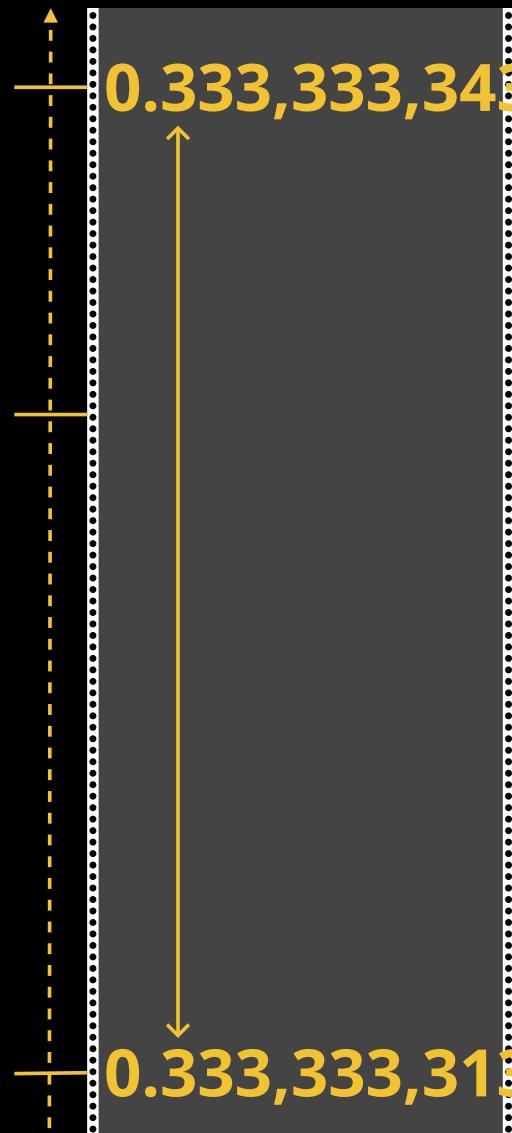


$$3,333,333,333,333,333,3 = 2^{-25} \times (11,184,810 + \frac{2}{3})$$

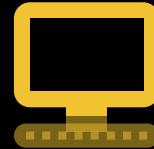


$$0.333,333,313,465,118,408,203,125,0 = 2^{-25} \times 11,184,810$$

$\frac{1}{3}$



$$0.333,333,343,267,440,795,898,437,5 = 2^{-25} \times 11,184,811$$

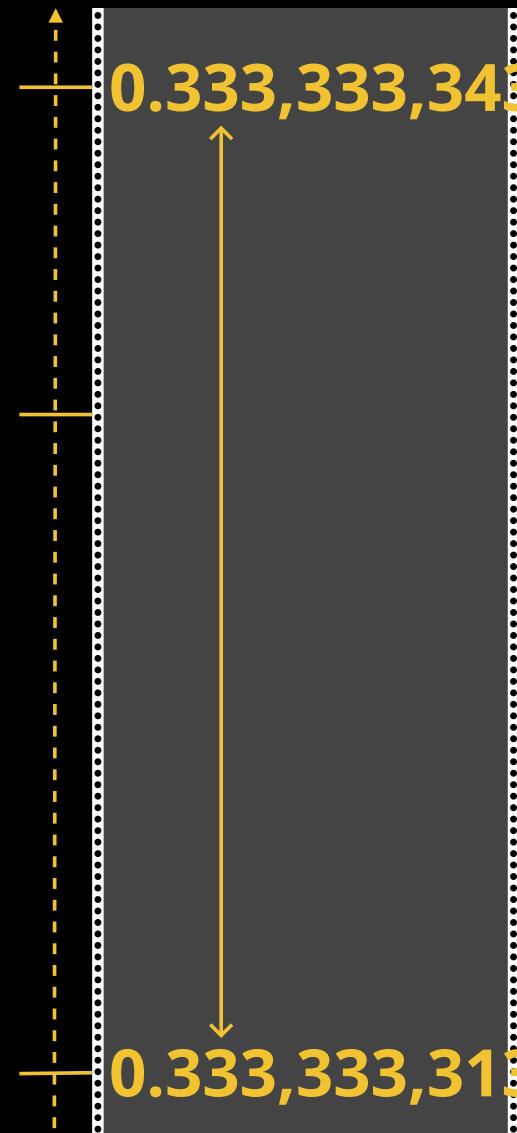


$$3,333,333,333,333,333,3 = 2^{-25} \times (11,184,810 + \frac{2}{3})$$

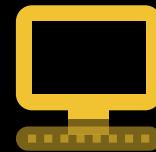


$$0.333,333,313,465,118,408,203,125,0 = 2^{-25} \times 11,184,810$$

$\frac{1}{3}$



$$0.333,333,343,267,440,795,898,437,5 = 2^{-25} \times 11,184,811$$

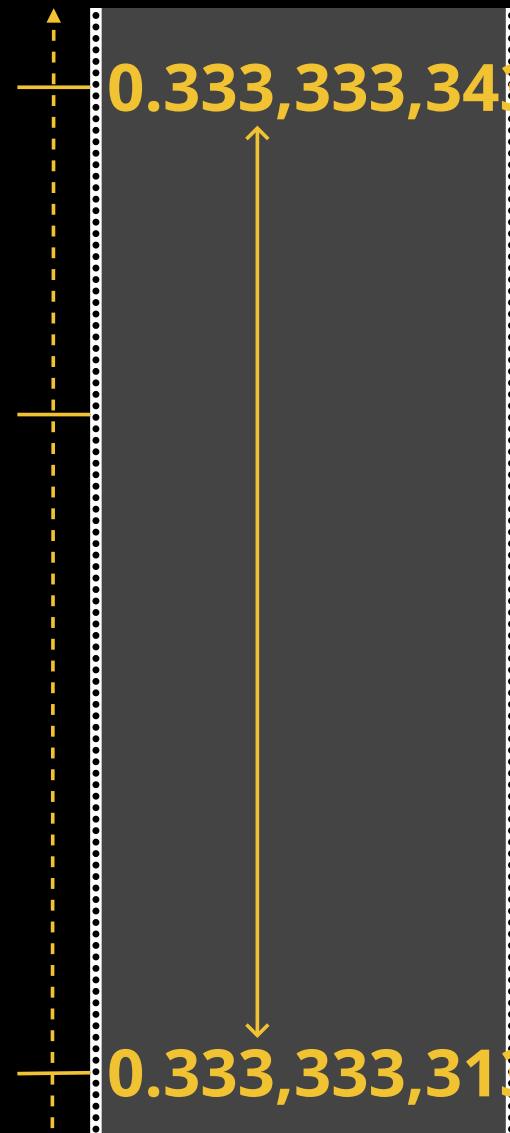


$$3,333,333,333,333,333,3 = 2^{-25} \times (11,184,810 + \frac{2}{3})$$

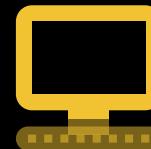


$$0.333,333,313,465,118,408,203,125,0 = 2^{-25} \times 11,184,810$$

$\frac{1}{3}$



$$0.333,333,343,267,440,795,898,437,5 = 2^{-25} \times 11,184,811$$

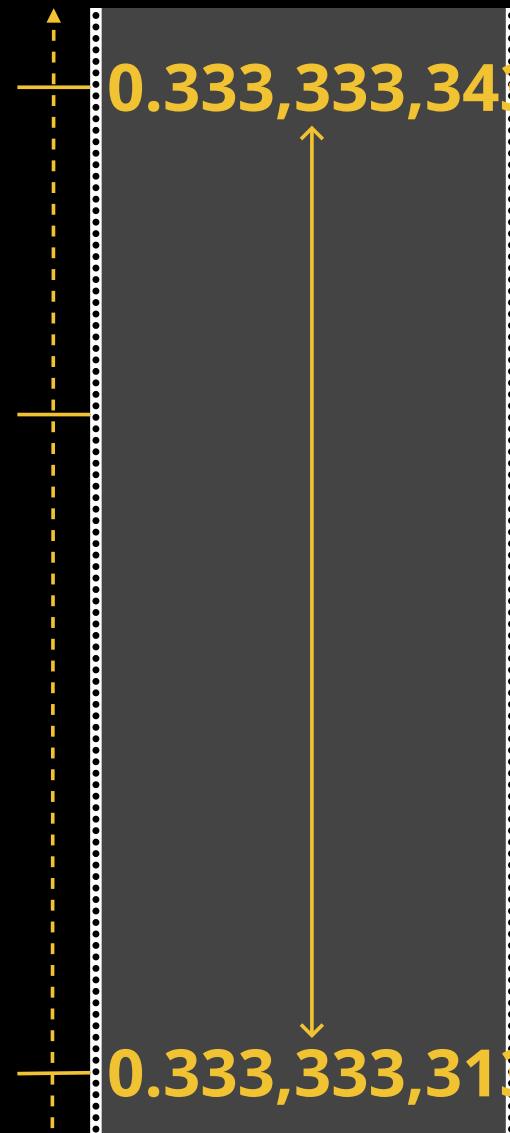


$$3,333,333,333,333,333,3 = 2^{-25} \times (11,184,810 + \frac{2}{3})$$

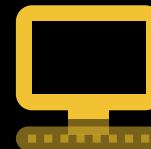


$$0.333,333,313,465,118,408,203,125,0 = 2^{-25} \times 11,184,810$$

$\frac{1}{3}$



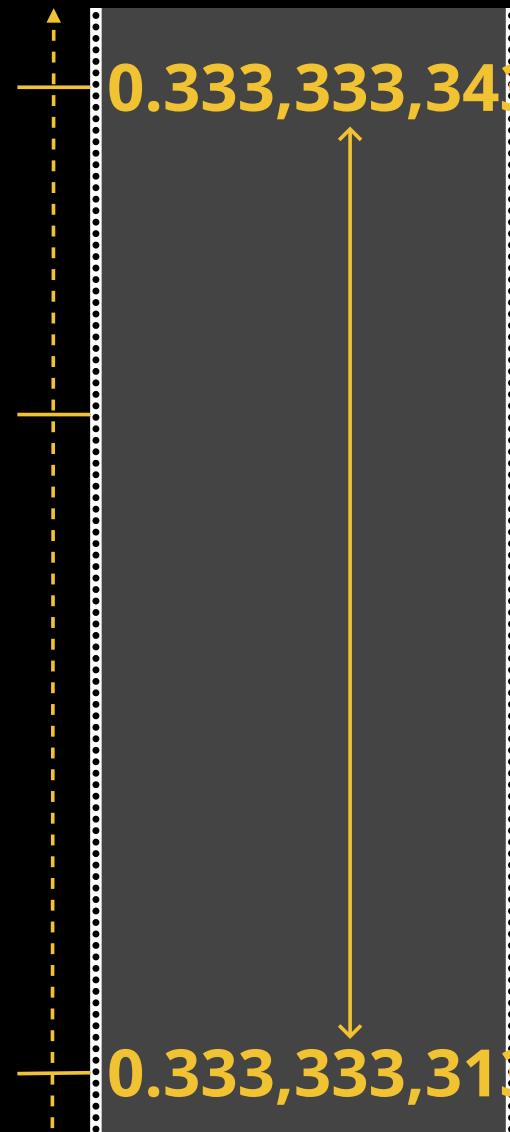
$$0.333,333,343,267,440,795,898,437,5 = 2^{-25} \times 11,184,811$$



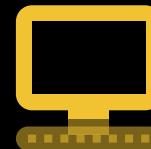
$$3,333,333,333,333,333,3 = 2^{-25} \times (11,184,810 + \frac{2}{3})$$



$$0.333,333,313,465,118,408,203,125,0 = 2^{-25} \times 11,184,810$$

$\frac{1}{3}$ 

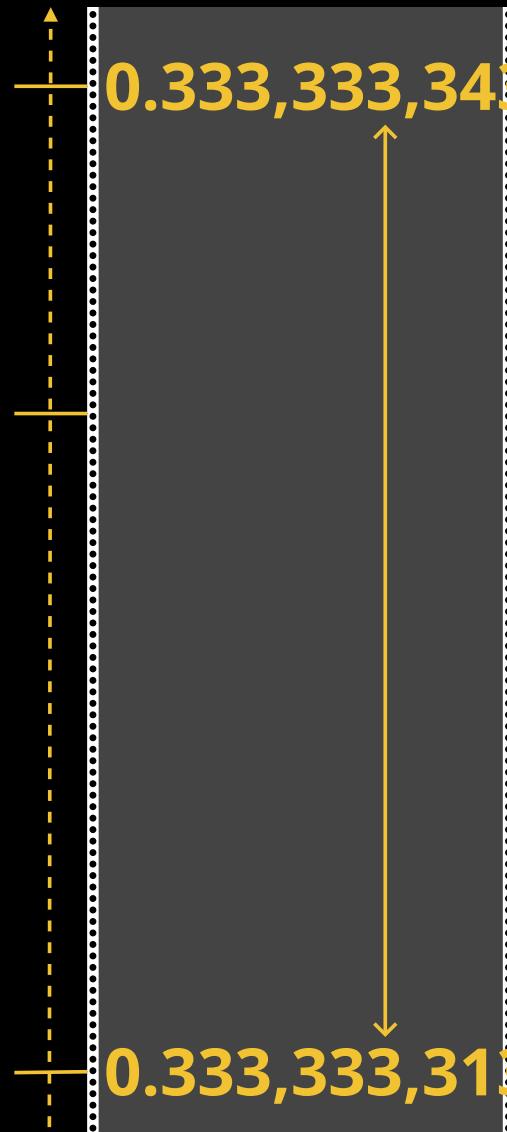
$$0.333,333,343,267,440,795,898,437,5 = 2^{-25} \times 11,184,811$$



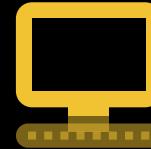
$$3,333,333,333,333,333,3 = 2^{-25} \times (11,184,810 + \frac{2}{3})$$



$$0.333,333,313,465,118,408,203,125,0 = 2^{-25} \times 11,184,810$$

$\frac{1}{3}$ 

$$0.333,333,343,267,440,795,898,437,5 = 2^{-25} \times 11,184,811$$

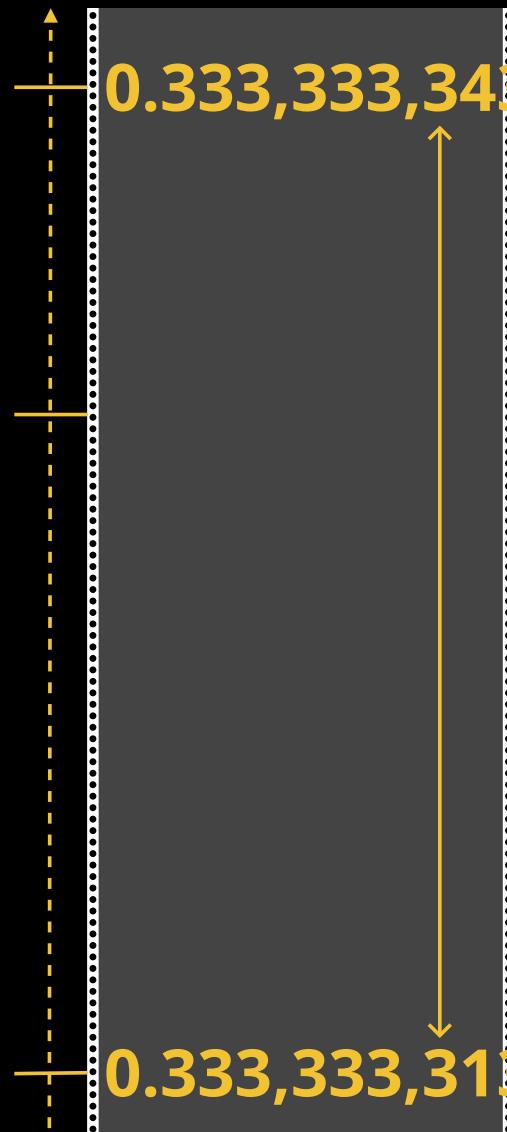


$$3,333,333,333,333,333,3 = 2^{-25} \times (11,184,810 + \frac{2}{3})$$

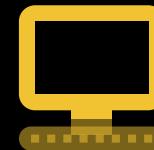


$$0.333,333,313,465,118,408,203,125,0 = 2^{-25} \times 11,184,810$$

$\frac{1}{3}$



$$0.333,333,343,267,440,795,898,437,5 = 2^{-25} \times 11,184,811$$



$$3,333,333,333,333,333,3 = 2^{-25} \times (11,184,810 + \frac{2}{3})$$

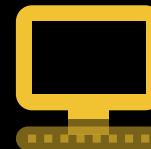


$$0.333,333,313,465,118,408,203,125,0 = 2^{-25} \times 11,184,810$$

$\frac{1}{3}$



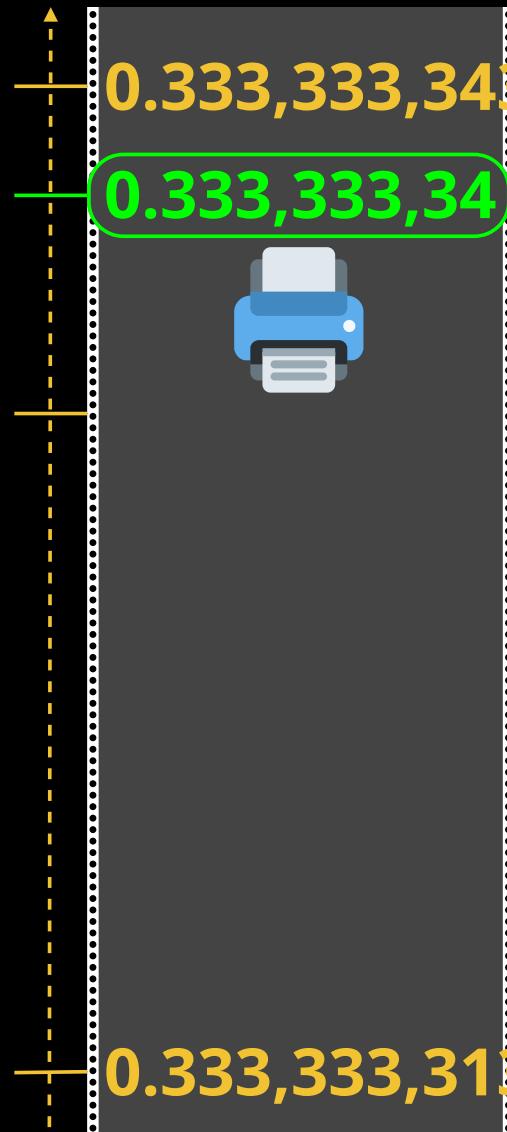
$$0.333,333,343,267,440,795,898,437,5 = 2^{-25} \times 11,184,811$$



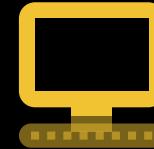
$$0.333,333,333,333,333,333,3 = 2^{-25} \times (11,184,810 + \frac{2}{3})$$

$$0.333,333,313,465,118,408,203,125,0 = 2^{-25} \times 11,184,810$$

$\frac{1}{3}$



$$0.333,333,34 = 2^{-25} \times 11,184,811$$



$$3,333,333,333,333,333,3 = 2^{-25} \times (11,184,810 + \frac{2}{3})$$



$$0.333,333,31 = 2^{-25} \times 11,184,810$$

# Jorge Luis Borges, The Book of Imaginary Beings

---



*"There is something in the dragon's image  
that fits man's imagination."*

# Dragon's Den

---



Copyright: BBC

# Dragon's Den

---



Copyright: BBC



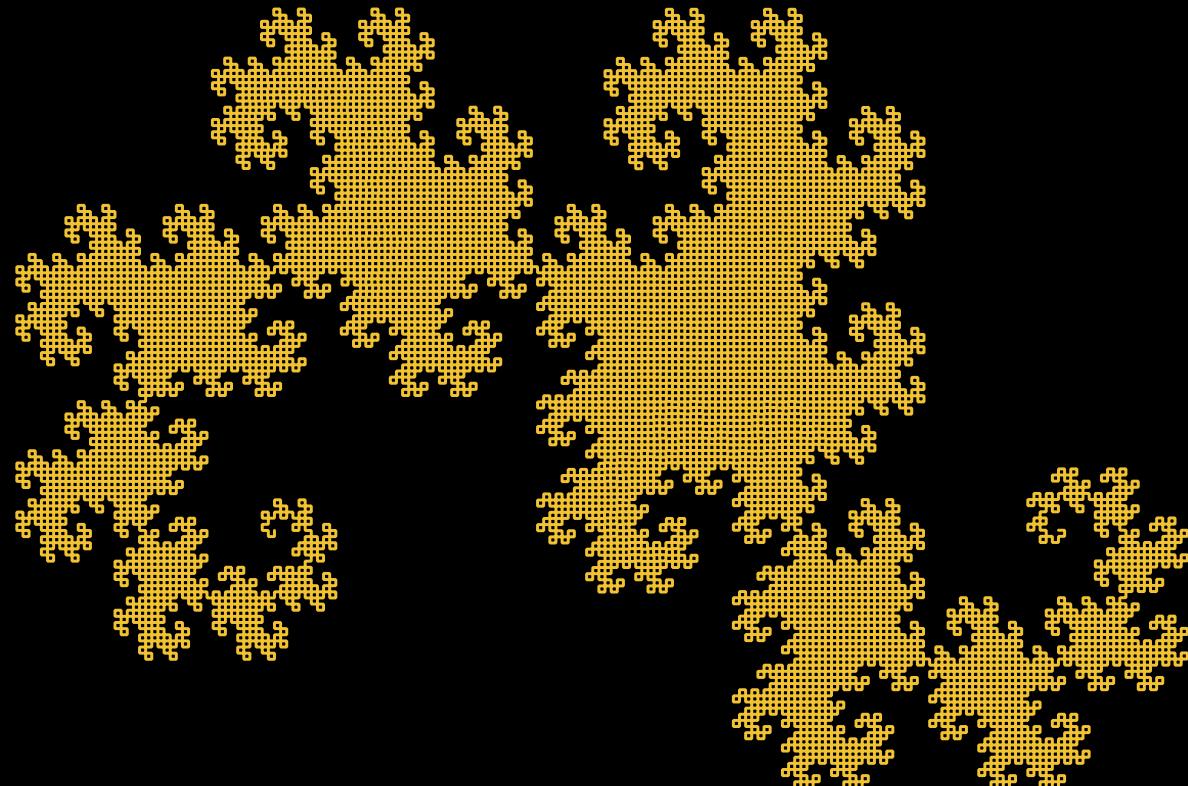
Copyright: ABC

# Another dragon's den

---

1990 - *Dragon*

Guy L. Steele, Jon L. White



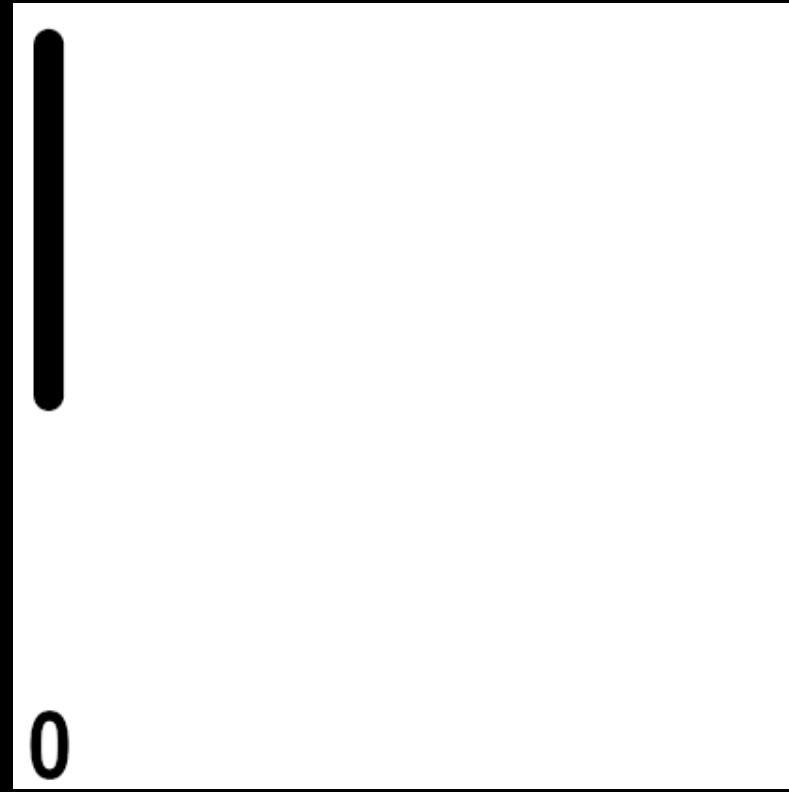
Heighway Dragon Curve  
Prokofiev, CC BY-SA 4.0, via Wikimedia Commons

# Another dragon's den

---

1990 - *Dragon*

Guy L. Steele, Jon L. White



Heighway Dragon Curve (orders 0 to 15)  
Public domain

# Another dragon's den

---

1990 - *Dragon*

Guy L. Steele, Jon L. White

1996

Robert G. Burger and R. Kent Dybvig

# Another dragon's den

---

1990 - *Dragon*

Guy L. Steele, Jon L. White

1996

Robert G. Burger and R. Kent Dyvig

2010 - *Grisù*

Florian Loitsch



Copyright: MONDO TV France S.A.

# Another dragon's den

---

1990 - <i>Dragon</i>	Guy L. Steele, Jon L. White
1996	Robert G. Burger and R. Kent Dybvig
2010 - <i>Grisù</i>	Florian Loitsch
2013	Aubrey Jaffer

# Another dragon's den

---

1990 - *Dragon*

Guy L. Steele, Jon L. White

1996

Robert G. Burger and R. Kent Dyvig

2010 - *Grisù*

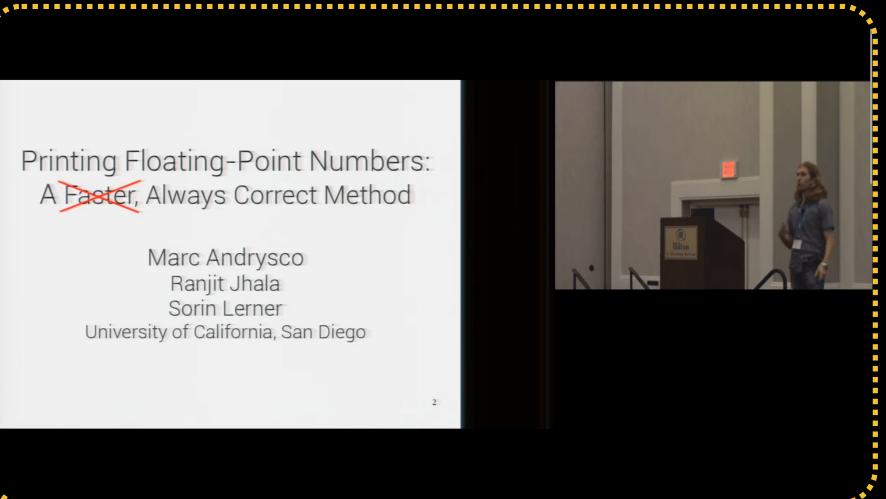
Florian Loitsch

2013

Aubrey Jaffer

2016 - *Errol*

Marc Andryesco, Ranjit Jhala, Sorin Lerner



# Another dragon's den

---

1990 - *Dragon*

Guy L. Steele, Jon L. White

1996

Robert G. Burger and R. Kent Dybvig

2010 - *Grisù*

Florian Loitsch

2013

Aubrey Jaffer

2016 - *Errol*

Marc Andryesco, Ranjit Jhala, Sorin Lerner

2018 - *Ryū*

Ulf Adams



Ryū - Fast Float-to-String Conversion  
Ulf Adams

Ryū: Fast Float-to-String Conversion  
Ulf Adams <ulfjack@google.com>

acm Association for Computing Machinery

This screenshot shows a presentation slide with a blue header and a green footer. The title is 'Ryū - Fast Float-to-String Conversion' and the author is 'Ulf Adams'. The slide content is 'Ryū: Fast Float-to-String Conversion' and the author's email 'Ulf Adams <ulfjack@google.com>'. The footer includes the ACM logo and the text 'Association for Computing Machinery'.

36 Ryu Techniques

- Ulf Adams' magic is still beyond my understanding
  - Read his paper and code, watch his talk
- Wide multiplications (64x128 for Ryu, 64x192 for Ryu Printf) followed by shifts
  - Multiplying by constants stored in large tables
  - Adams proved that arbitrary precision is not necessary
- Produces integers (e.g. 1729), writes "17.29"
  - Core algorithm is so fast, this step is relatively costly!
- Only integer operations; cold FPU transistors

Floating-Point <charconv>:  
Making Your Code 10x  
Faster With C++17's  
Final Boss

Video Sponsorship Provided By:  
ansatz

Cppcon | 2019  
The C++ Conference  
cppcon.org

Stephan T. Lavavej

This screenshot shows a presentation slide titled 'Ryu Techniques' with a slide number '36'. The slide content lists several bullet points about Ryu's techniques, such as wide multiplications, integer production, and the core algorithm's speed. It also mentions Ulf Adams' work and the final boss of floating-point conversion. The slide is from the Cppcon 2019 conference. A video sponsorship for 'ansatz' is mentioned at the bottom.

# Another dragon's den

---

1990 - <i>Dragon</i>	Guy L. Steele, Jon L. White
1996	Robert G. Burger and R. Kent Dybvig
2010 - <i>Grisù</i>	Florian Loitsch
2013	Aubrey Jaffer
2016 - <i>Errol</i>	Marc Andryesco, Ranjit Jhala, Sorin Lerner
2018 - <i>Ryū</i>	Ulf Adams
2020 - <i>Schubfach</i>	Raffaello Giulietti

# Another dragon's den

---

1990 - <i>Dragon</i>	Guy L. Steele, Jon L. White
1996	Robert G. Burger and R. Kent Dybvig
2010 - <i>Grisù</i>	Florian Loitsch
2013	Aubrey Jaffer
2016 - <i>Errol</i>	Marc Andryesco, Ranjit Jhala, Sorin Lerner
2018 - <i>Ryū</i>	Ulf Adams
2020 - <i>Schubfach</i>	Raffaello Giulietti
2020 - <i>Grisù-Exact</i>	Junekey Jeon

# Another dragon's den

---

1990 - <i>Dragon</i>	Guy L. Steele, Jon L. White
1996	Robert G. Burger and R. Kent Dyvig
2010 - <i>Grisù</i>	Florian Loitsch
2013	Aubrey Jaffer
2016 - <i>Errol</i>	Marc Andryesco, Ranjit Jhala, Sorin Lerner
2018 - <i>Ryū</i>	Ulf Adams
2020 - <i>Schubfach</i>	Raffaello Giulietti
2020 - <i>Grisù-Exact</i>	Junekey Jeon
2022 - <i>Dragonbox</i>	Junekey Jeon



# Three steps

---

# Three steps

---

Binary32		
±	biased exponent	mantissa
1	8	23
0 1 0 0 0 0 0 1 0 0 1 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		

$+2^{00010100} \times 10101100000000000000000000000000$

$\oplus^{\text{20}}$  x [ 1 1 , 2 7 2 , 1 9 2 ]

decode representation

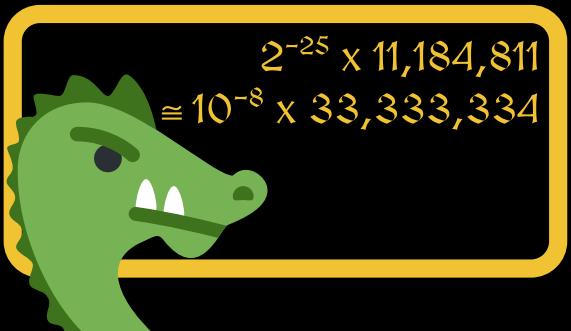
# Three steps

---

Binary32		
±	biased exponent	mantissa
1	8	23
0 1 0 0 0 0 0 1 0 0 1 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		

$+2^{00010100} \times 10101100000000000000000000000000$

$\oplus^{\text{20}} \times [1 \ 1, 2 \ 7 \ 2, 1 \ 9 \ 2]$



decode representation



convert binary to decimal

# Three steps

---



decode representation



convert binary to decimal



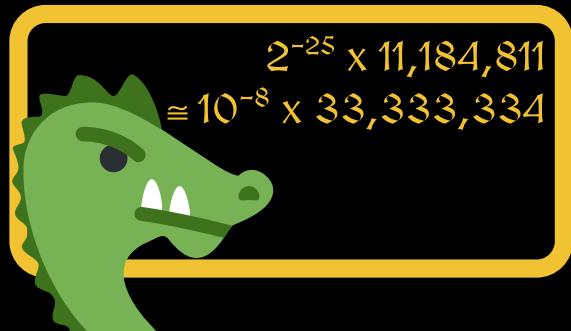
convert exponent and mantissa to string



# Dragon's problem

---

Given  $m \in \mathbb{N}$  and  $E \in \mathbb{Z}$ ,  
find  $n \in \mathbb{N}$  and  $F \in \mathbb{Z}$  such that  
 $n \times 10^F \cong m \times 2^E$ .



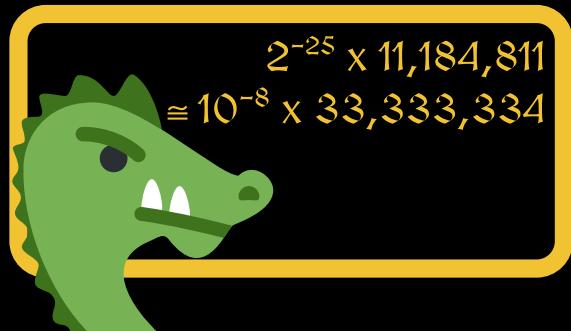
# Dragon's problem

---

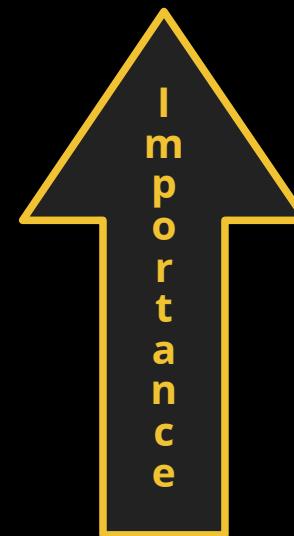
Given  $m \in \mathbb{N}$  and  $E \in \mathbb{Z}$ ,

find  $n \in \mathbb{N}$  and  $F \in \mathbb{Z}$  such that

$$n \times 10^F \cong m \times 2^E.$$

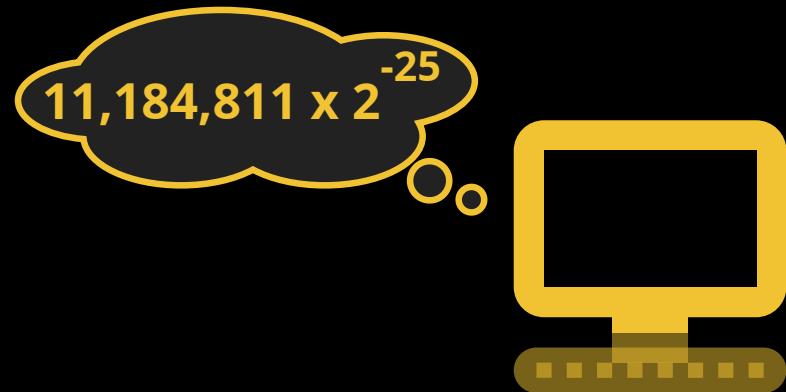


- No information loss
- As short as possible
- As close as possible
- Tiebreak rules



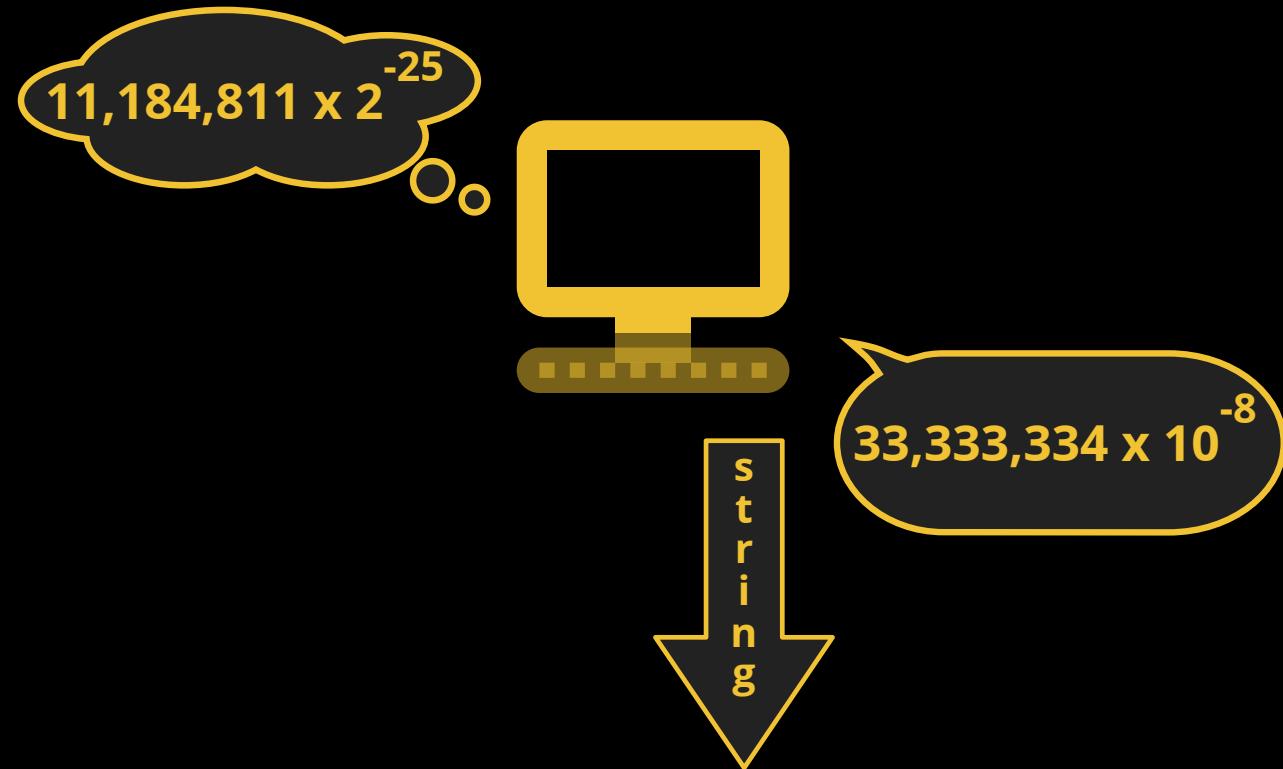
# No information loss

---



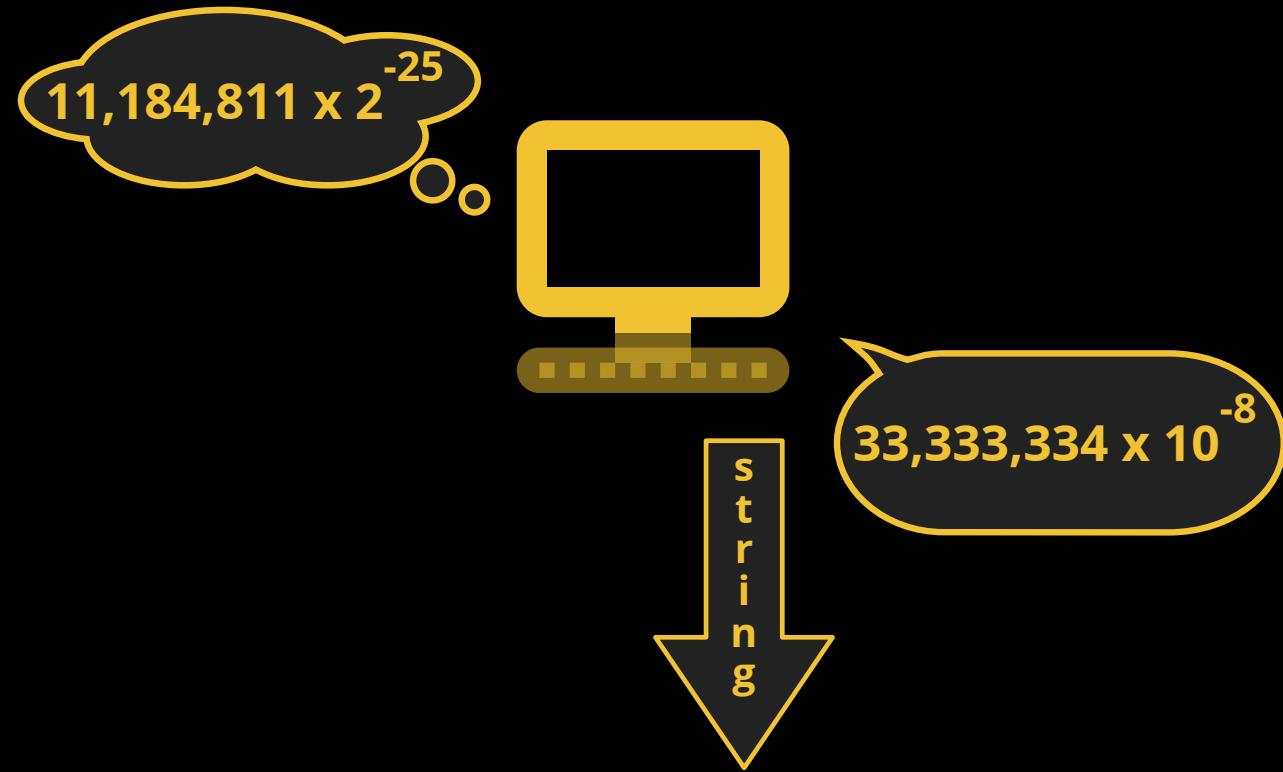
# No information loss

---



# No information loss

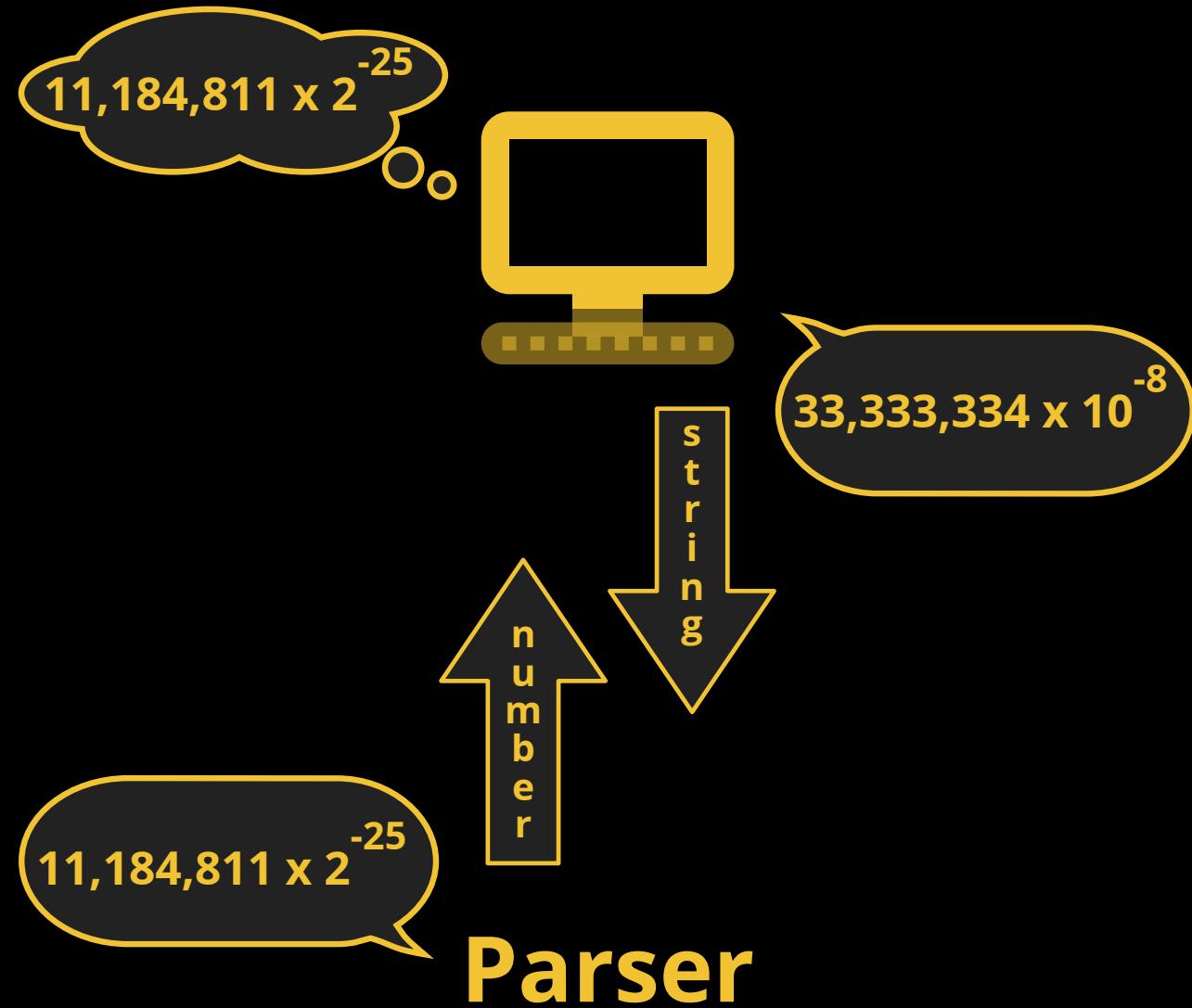
---



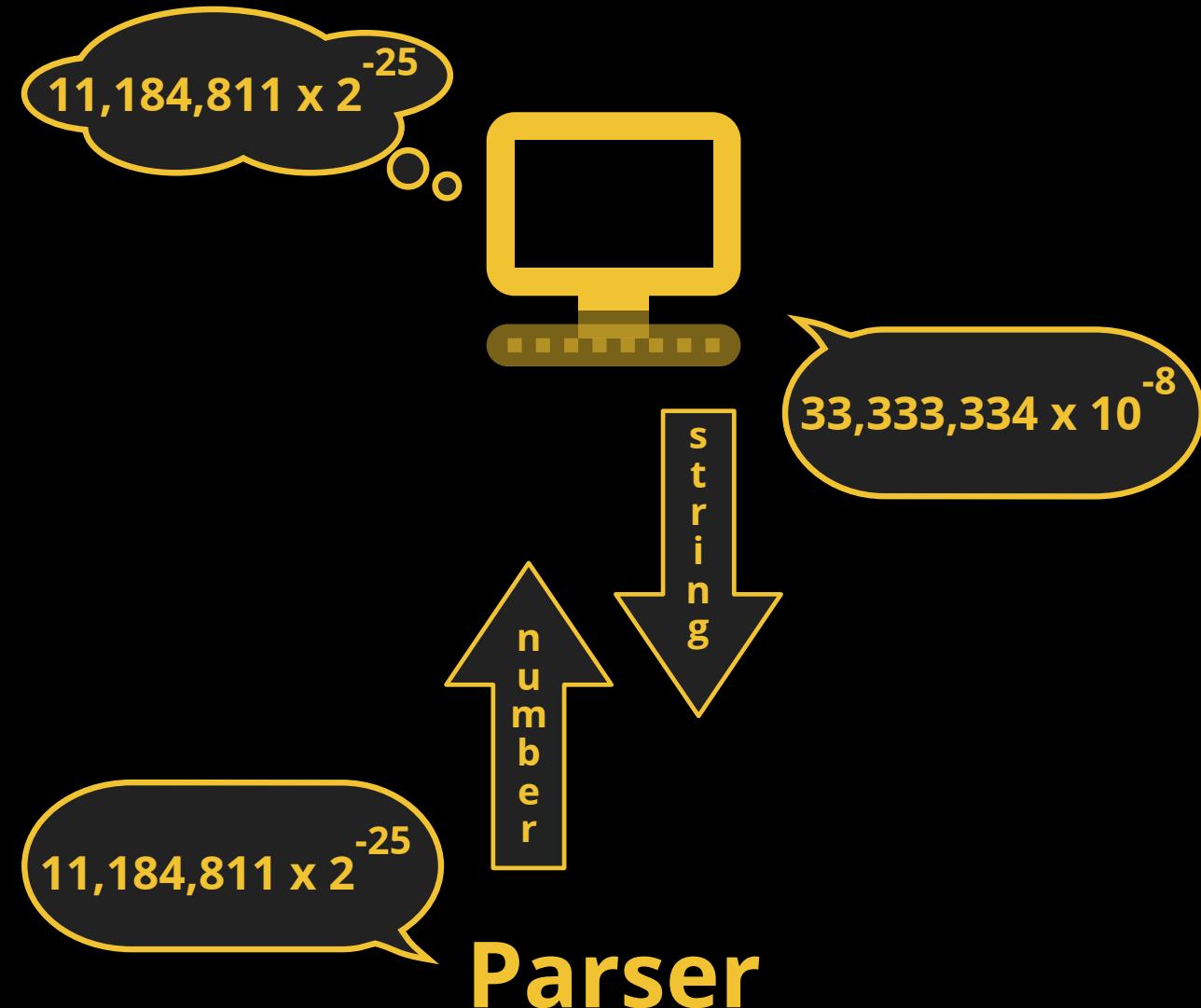
# Parser

# No information loss

---



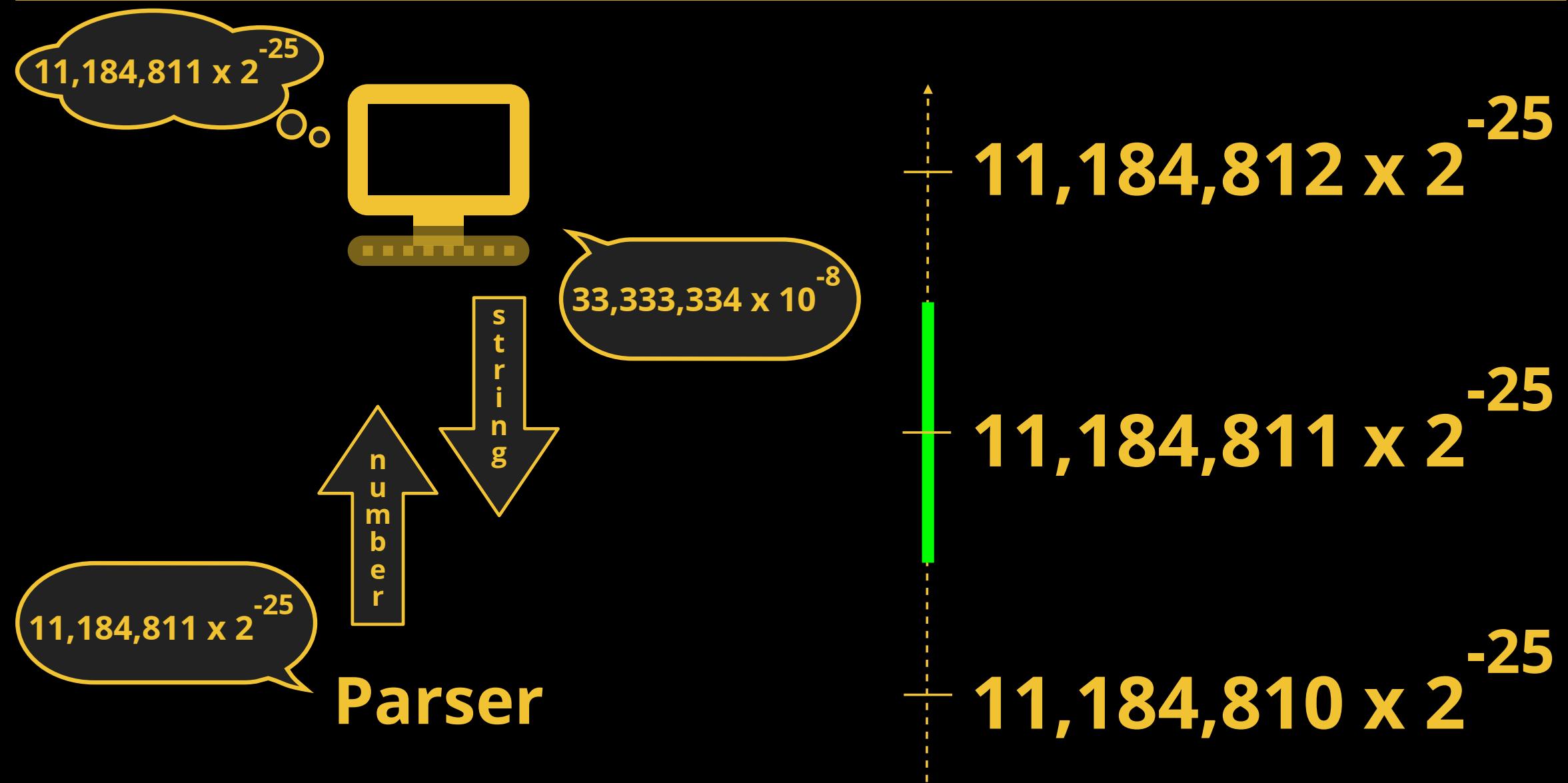
# No information loss



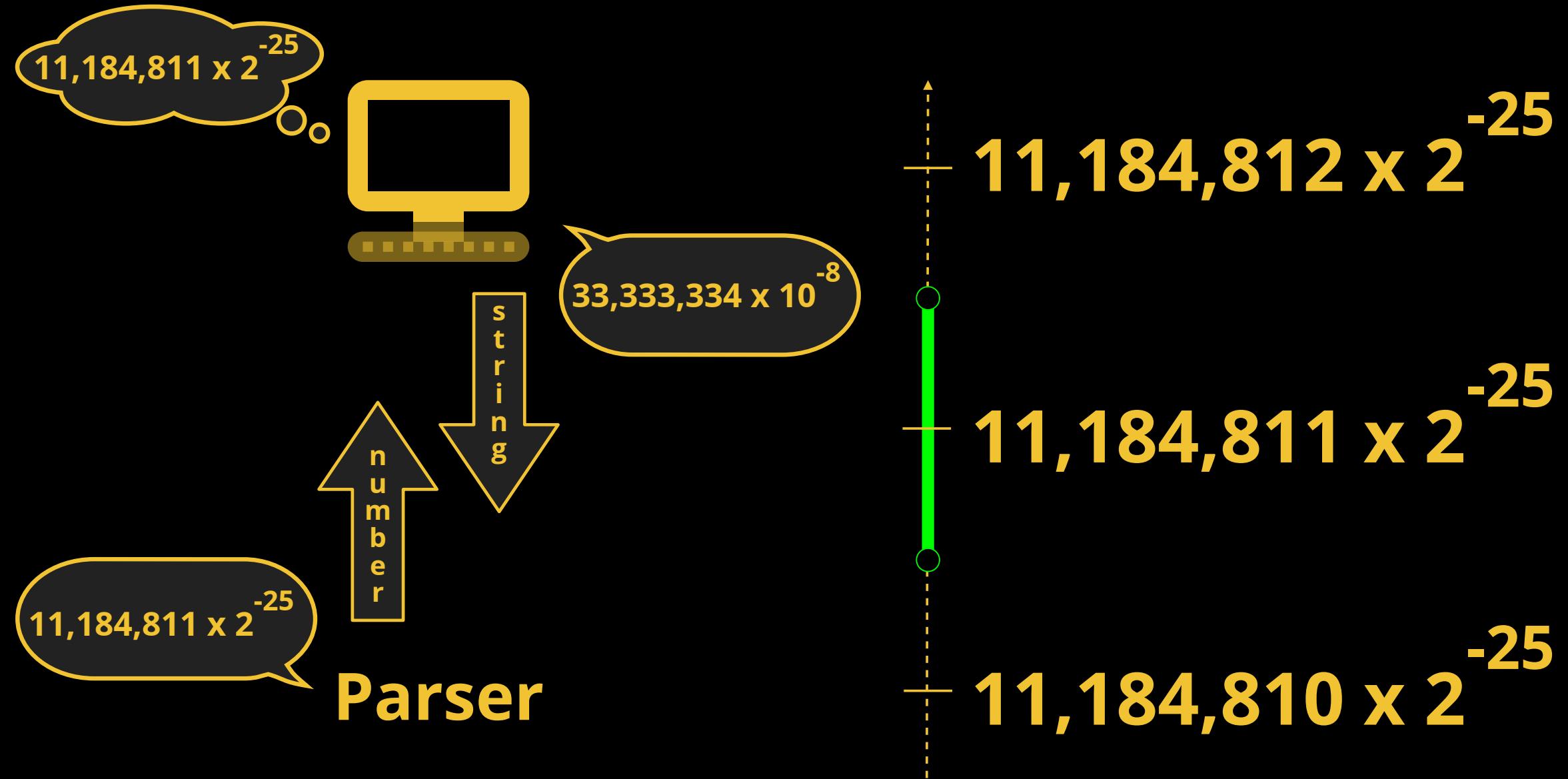
A vertical dashed line is positioned to the left of three floating-point numbers, each enclosed in a thought bubble:

- $11,184,812 \times 2^{-25}$
- $11,184,811 \times 2^{-25}$
- $11,184,810 \times 2^{-25}$

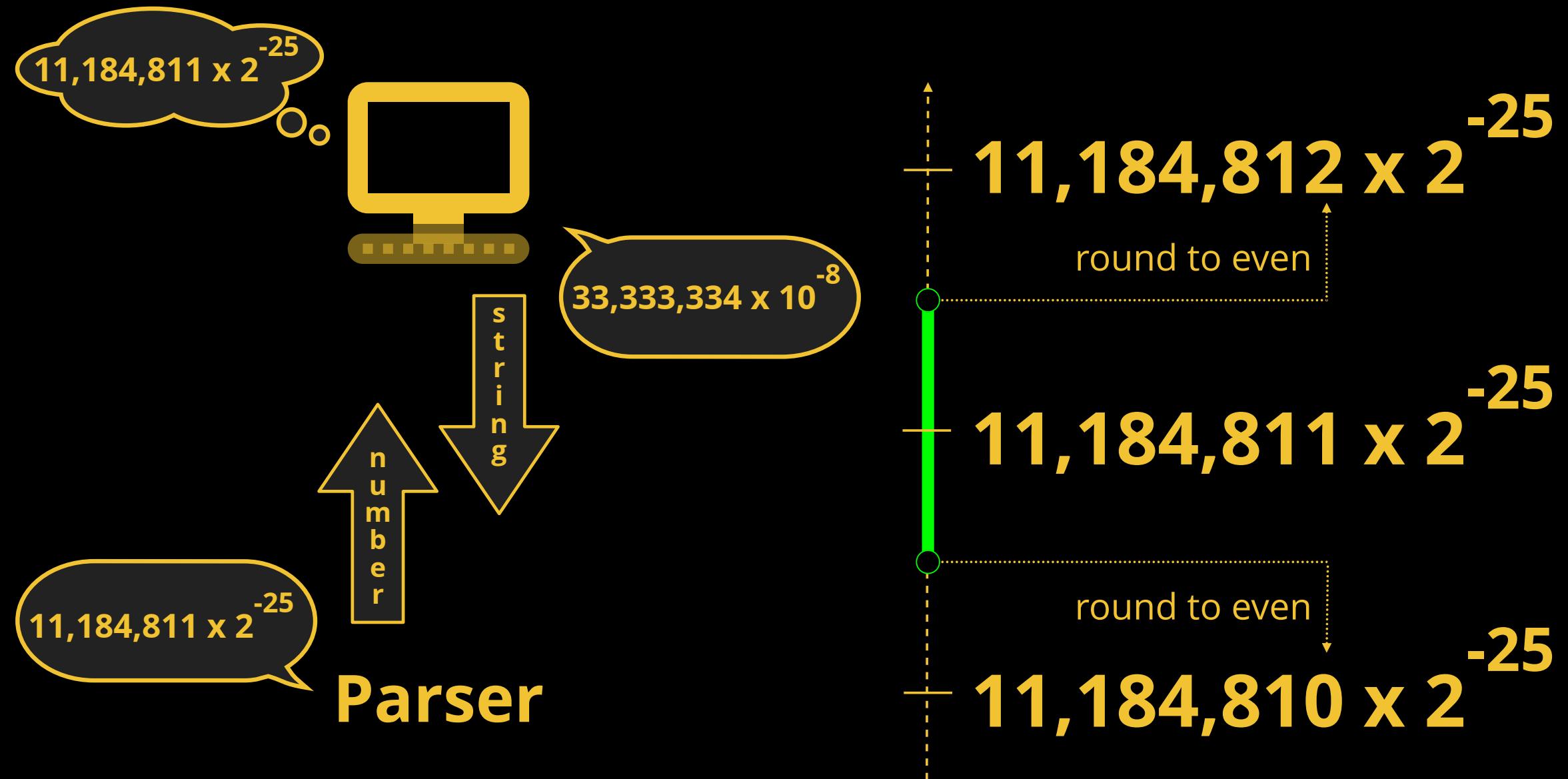
# No information loss



# No information loss



# No information loss



As short as possible

---

$$9,765,625 \times 2^{10} = 10 \times 10^9$$

$$= 1 \times 10^{10}$$

As short as possible

---

$$9,765,625 \times 2^{10} = 10 \times 10^9$$

$$= 1 \times 10^{10}$$

Only the mantissa matters.

As short as possible

---

$$13,421,773 \times 2^{-27} \approx 100 \times 10^{-3}$$

$$\approx 10 \times 10^{-2}$$

$$\approx 1 \times 10^{-1}$$

# As short as possible

---

$$13,421,773 \times 2^{-27} \approx 100 \times 10^{-3}$$

$$\approx 10 \times 10^{-2}$$

$$\approx 1 \times 10^{-1}$$

$$\approx 0.1$$

# As short as possible

---

$$13,421,773 \times 2^{-27} \approx 100 \times 10^{-3}$$

$$\approx 10 \times 10^{-2}$$

$$\approx 1 \times 10^{-1}$$

$$\approx 0.1$$

only representations of  
the form  $n \times 10^F$  matter



# As short as possible

---

$$13,421,773 \times 2^{-27} \approx 100 \times 10^{-3}$$

$$\approx 10 \times 10^{-2}$$

$$\approx 1 \times 10^{-1}$$

$\approx$



only representations of  
the form  $n \times 10^F$  matter



# As short as possible

---

$$13,421,773 \times 2^{-27} \approx 100 \times 10^{-3}$$

$$\approx 10 \times 10^{-2}$$

only representations of  
the form  $n \times 10^F$  matter

$$\approx 1 \times 10^{-1}$$

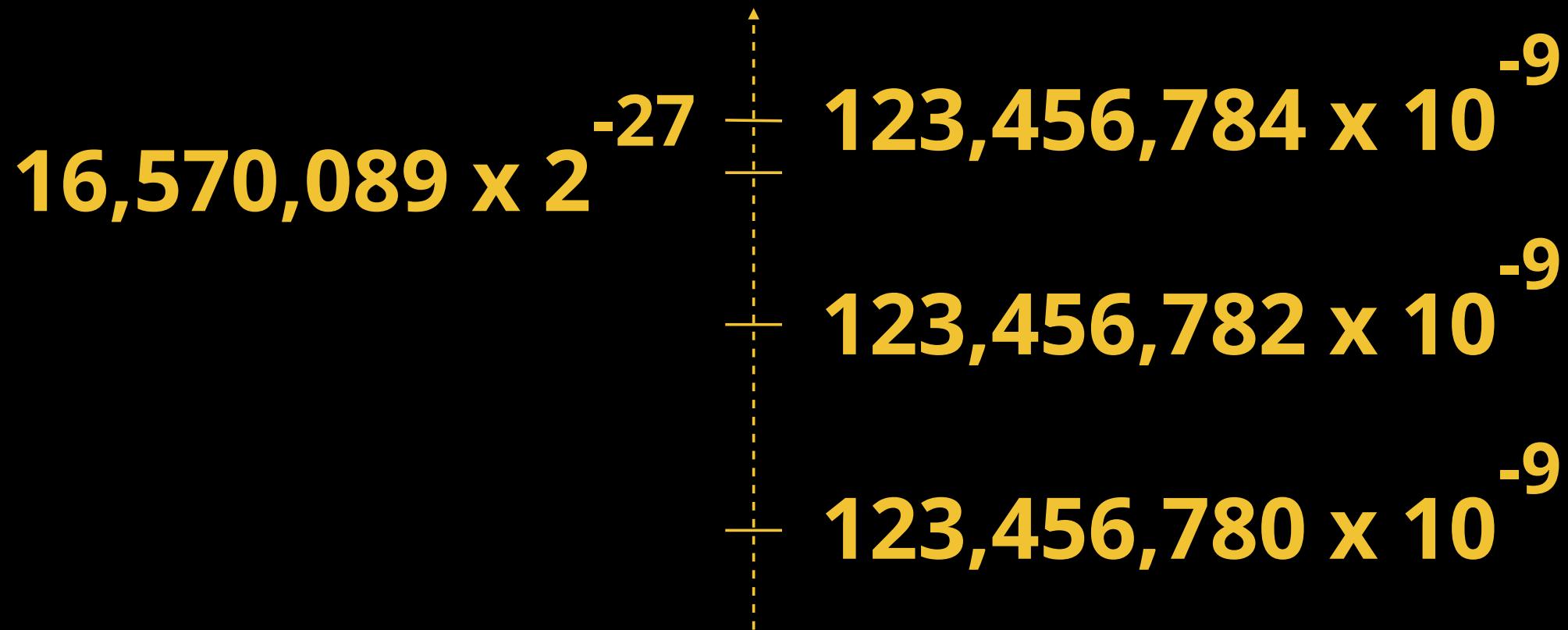
the larger  $F$ ,  
the shorter  $n$

≈



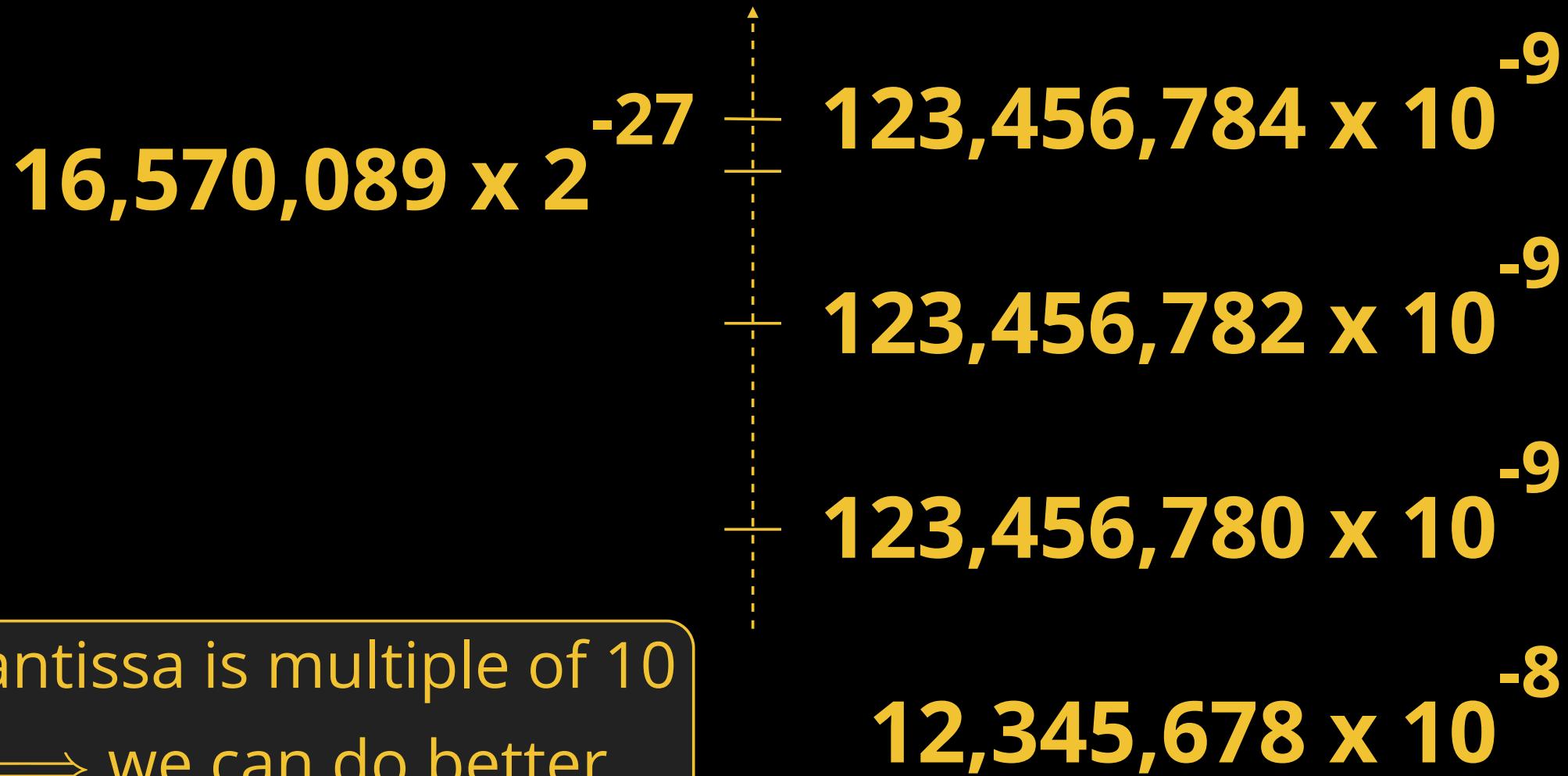
# As short as possible

---



# As short as possible

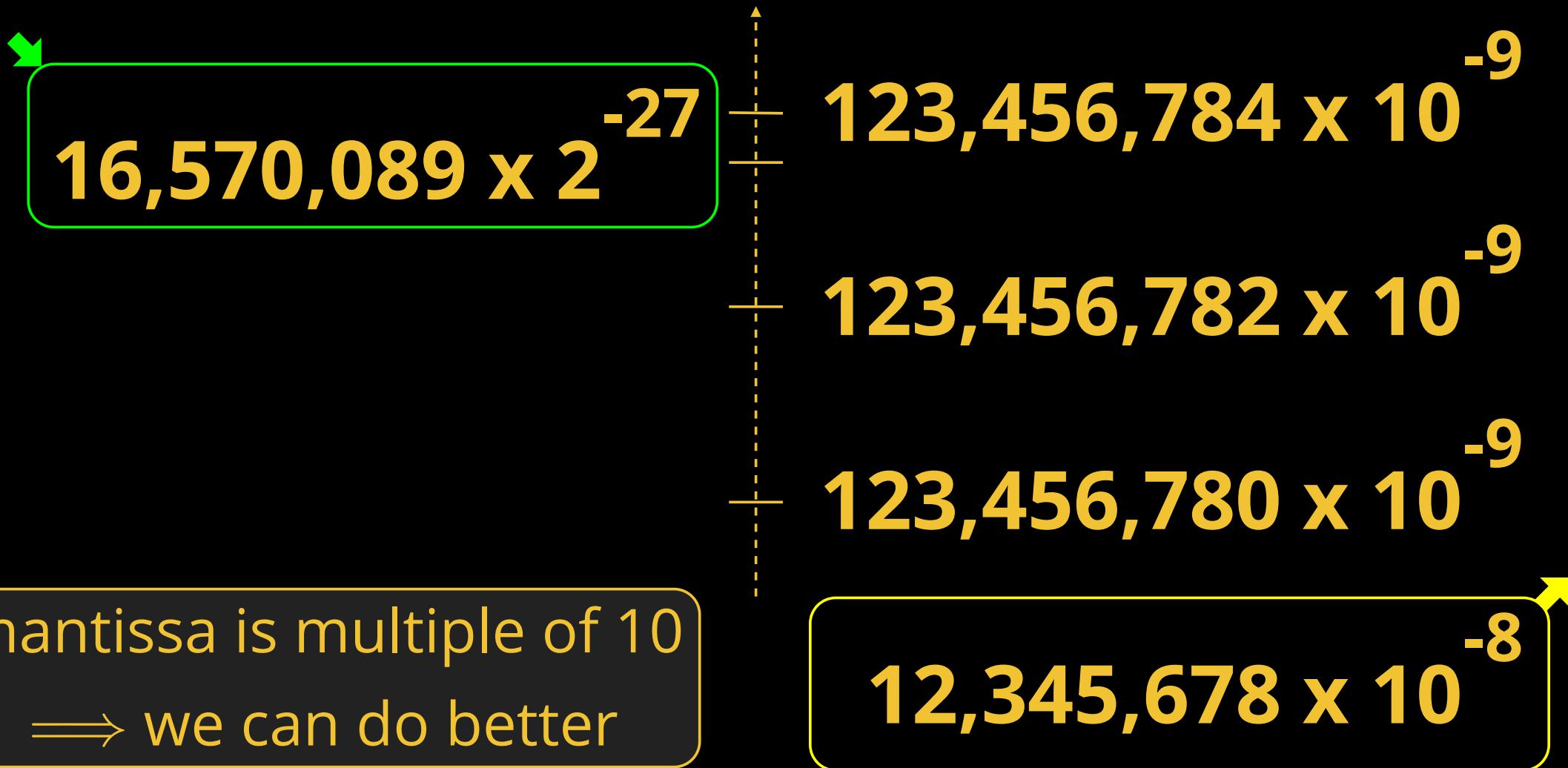
---



mantissa is multiple of 10  
⇒ we can do better

# As short as possible

---



# As short as possible

---

$10,066,330 \times 2^{-25}$



$30,000,002 \times 10^{-8}$

$30,000,001 \times 10^{-8}$

$30,000,000 \times 10^{-8}$

# As short as possible

---

$10,066,330 \times 2^{-25}$



$30,000,002 \times 10^{-8}$

$30,000,001 \times 10^{-8}$

$30,000,000 \times 10^{-8}$

$3 \times 10^{-1}$

mantissa is multiple of 10  
 $\Rightarrow$  we can do better

## As short as possible

---



$10,066,330 \times 2^{-25}$



$30,000,002 \times 10^{-8}$

$30,000,001 \times 10^{-8}$

$30,000,000 \times 10^{-8}$

mantissa is multiple of 10  
 $\Rightarrow$  we can do better

$3 \times 10^{-1}$

# Pythagoras

---



*“Leave the roads;  
take the trails.”*

# Guaraní people

---



# Guaraní mythology

---



# Guaraní mythology

---

1 God Tupã and goddess Arasy created the first humans, Rupave and Sypave.



# Guaraní mythology

---

1 God Tupã and goddess Arasy created the first humans, Rupave and Sypave.



2 Tupã created the spirits of good, Angatupyry, and evil, Taú.



# Guaraní mythology

---

1 God Tupã and goddess Arasy created the first humans, Rupave and Sypave.

2 Tupã created the spirits of good, Angatupyry, and evil, Taú.

3 Marangatú was the second son of Rupave and Sypave.



# Guaraní mythology

---

1 God Tupã and goddess Arasy created the first humans, Rupave and Sypave.

2 Tupã created the spirits of good, Angatupyry, and evil, Taú.

3 Marangatú was the second son of Rupave and Sypave.

4 Taú felt in love with Keraná, Marangatú's beautiful daughter.



# Guaraní mythology

1 God Tupã and goddess Arasy created the first humans, Rupave and Sypave.

2 Tupã created the spirits of good, Angatupyry, and evil, Taú.

3 Marangatú was the second son of Rupave and Sypave.

4 Taú felt in love with Keraná, Marangatú's beautiful daughter.

5 Disguised as a handsome young man, Taú defeated Angatupyry and captured Keraná.



# Guaraní mythology

1 God Tupã and goddess Arasy created the first humans, Rupave and Sypave.

2 Tupã created the spirits of good, Angatupyry, and evil, Taú.

3 Marangatú was the second son of Rupave and Sypave.

4 Taú felt in love with Keraná, Marangatú's beautiful daughter.

5 Disguised as a handsome young man, Taú defeated Angatupyry and captured Keraná.



6 Arasy, placed a curse upon Taú's seven sons, who were born as terrifying beasts.



# Guaraní mythology

1 God Tupã and goddess Arasy created the first humans, Rupave and Sypave.

2 Tupã created the spirits of good, Angatupyry, and evil, Taú.

3 Marangatú was the second son of Rupave and Sypave.

4 Taú felt in love with Keraná, Marangatú's beautiful daughter.

5 Disguised as a handsome young man, Taú defeated Angatupyry and captured Keraná.



6 Arasy, placed a curse upon Taú's seven sons, who were born as terrifying beasts.

7 Tejú Jaguá was the first legendary monster.

# Tejú Jaguá - I

---

$$n \times 10^F \cong m \times 2^E$$

# Tejú Jaguá - I

$$n \times 10^F \cong m \times 2^E$$

$$m \times 2^E$$

# Tejú Jaguá - I

---

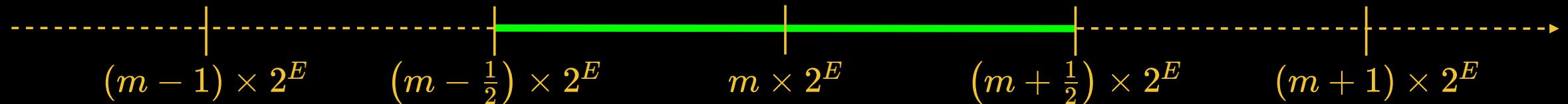
$$n \times 10^F \cong m \times 2^E$$

$$\overbrace{\qquad\qquad\qquad}^{(m-1) \times 2^E} \qquad \overbrace{\qquad\qquad\qquad}^{m \times 2^E} \qquad \overbrace{\qquad\qquad\qquad}^{(m+1) \times 2^E} \rightarrow$$

# Tejú Jaguá - I

---

$$n \times 10^F \cong m \times 2^E$$



# Tejú Jaguá - I

---

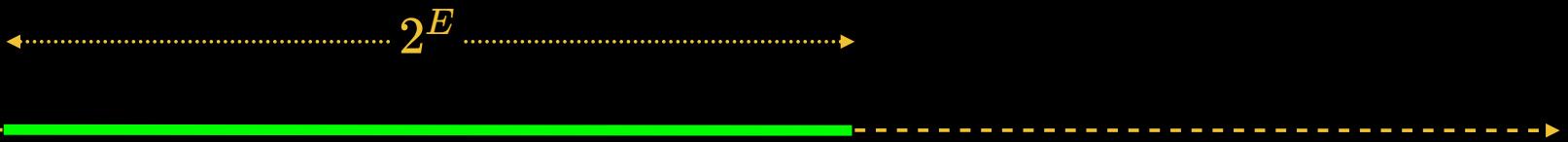
$$n \times 10^F \cong m \times 2^E$$

A horizontal dotted line with arrows at both ends spans the width of the green bar. Below it, five vertical dashed lines define segments of length  $(m + \frac{1}{2}) \times 2^E$ . The first segment is labeled  $(m - 1) \times 2^E$ , the second  $(m - \frac{1}{2}) \times 2^E$ , the third  $m \times 2^E$ , the fourth  $(m + \frac{1}{2}) \times 2^E$ , and the fifth  $(m + 1) \times 2^E$ .

# Tejú Jaguá - I

---

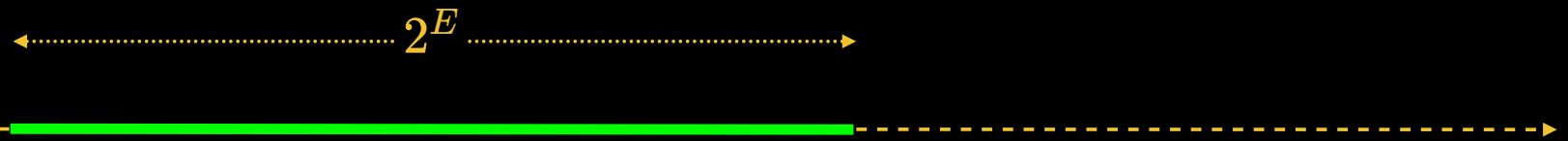
$$n \times 10^F \cong m \times 2^E$$



# Tejú Jaguá - I

---

$$n \times 10^F \cong m \times 2^E$$

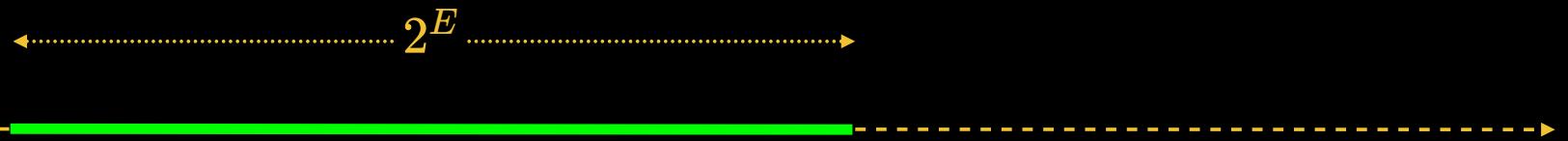


- The larger  $F$ , the shorter  $n$ .

# Tejú Jaguá - I

---

$$n \times 10^F \cong m \times 2^E$$

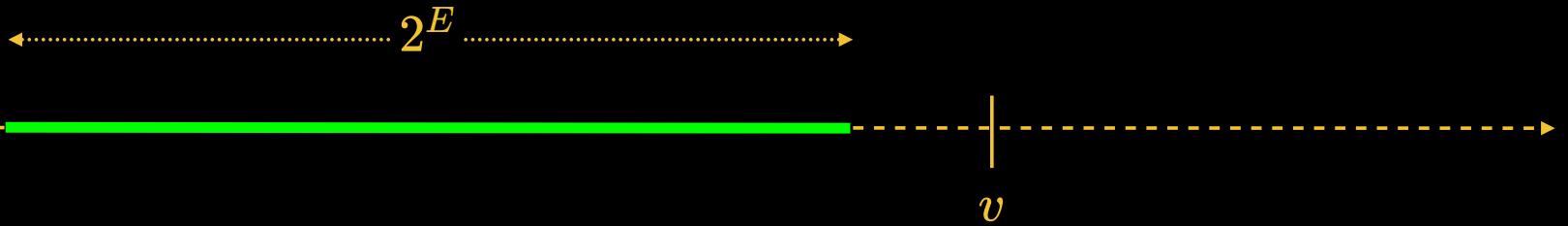


- The larger  $F$ , the shorter  $n$ .
- We wish to ensure that  $n \times 10^F$  falls inside the permissible interval.

# Tejú Jaguá - I

---

$$n \times 10^F \cong m \times 2^E$$

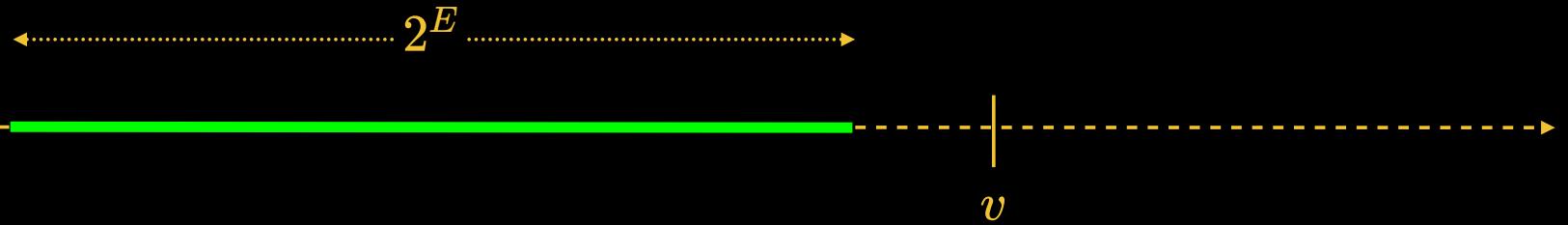


- The larger  $F$ , the shorter  $n$ .
- We wish to ensure that  $n \times 10^F$  falls inside the permissible interval.
- If we fail, then two **consecutive** numbers  $u$  and  $v$  of this form fall outside the permissible interval and thus,  $10^F = v - u > 2^E$ .

# Tejú Jaguá - I

---

$$n \times 10^F \cong m \times 2^E$$

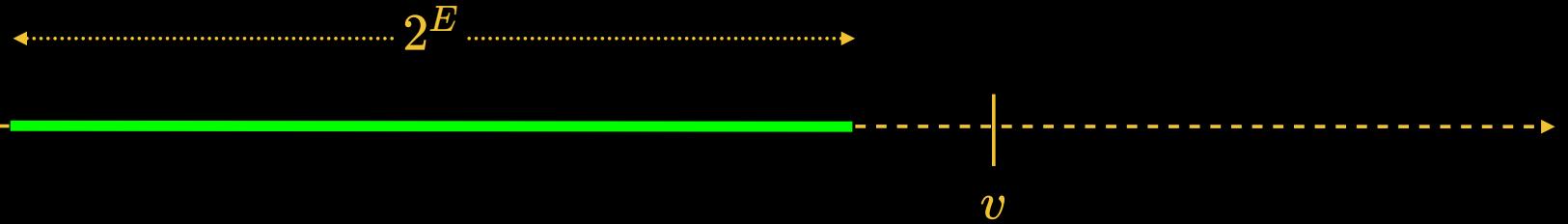


- The larger  $F$ , the shorter  $n$ .
- We wish to ensure that  $n \times 10^F$  falls inside the permissible interval.
- If we fail, then two **consecutive** numbers  $u$  and  $v$  of this form fall outside the permissible interval and thus,  $10^F = v - u > 2^E$ .
- We set  $F$  to be the largest integer such that  $10^F \leq 2^E$ , i.e.,

# Tejú Jaguá - I

---

$$n \times 10^F \cong m \times 2^E$$



- The larger  $F$ , the shorter  $n$ .
- We wish to ensure that  $n \times 10^F$  falls inside the permissible interval.
- If we fail, then two **consecutive** numbers  $u$  and  $v$  of this form fall outside the permissible interval and thus,  $10^F = v - u > 2^E$ .
- We set  $F$  to be the largest integer such that  $10^F \leq 2^E$ , i.e.,

$$10^F \leq 2^E < 10^{F+1} \iff F \leq E \times \log(2) < F + 1 \iff F = \lfloor E \times \log(2) \rfloor$$

# Tejú Jaguá - II

---

$$\left(m - \frac{1}{2}\right) \times 2^E$$

$$m \times 2^E$$

$$\left(m + \frac{1}{2}\right) \times 2^E$$



# Tejú Jaguá - II

---

$$(2m - 1) \times 2^{E-1}$$

$$2m \times 2^{E-1}$$

$$(2m + 1) \times 2^{E-1}$$



# Tejú Jaguá - II

---

$$(2m - 1) \times 2^{E-1}$$

$$2m \times 2^{E-1}$$

$$(2m + 1) \times 2^{E-1}$$



★ Points of the form  $n \times 10^F$

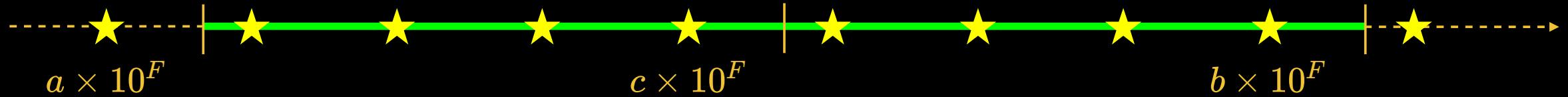
# Tejú Jaguá - II

---

$$(2m - 1) \times 2^{E-1}$$

$$2m \times 2^{E-1}$$

$$(2m + 1) \times 2^{E-1}$$



★ Points of the form  $n \times 10^F$

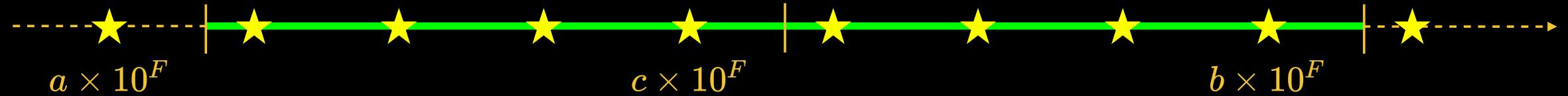
# Tejú Jaguá - II

---

$$(2m - 1) \times 2^{E-1}$$

$$2m \times 2^{E-1}$$

$$(2m + 1) \times 2^{E-1}$$

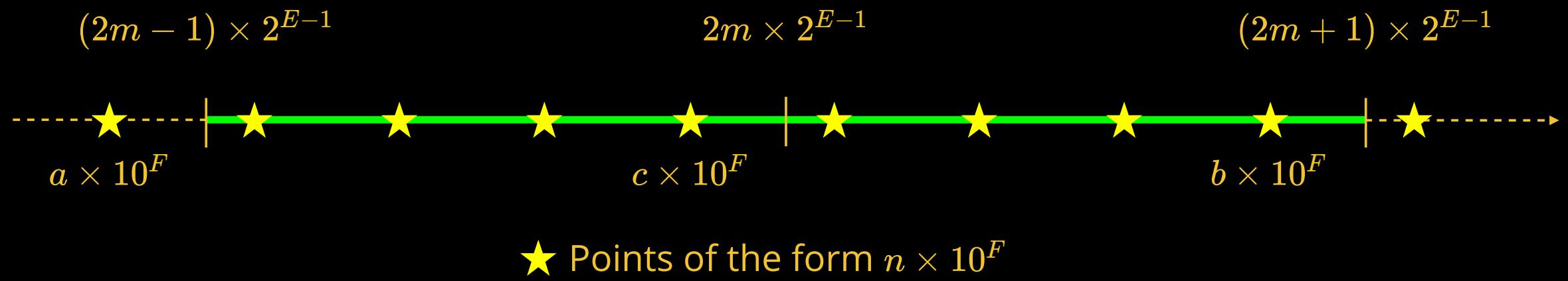


★ Points of the form  $n \times 10^F$

$$a \times 10^F \leq (2m - 1) \times 2^{E-1} < (a + 1) \times 10^F$$

# Tejú Jaguá - II

---

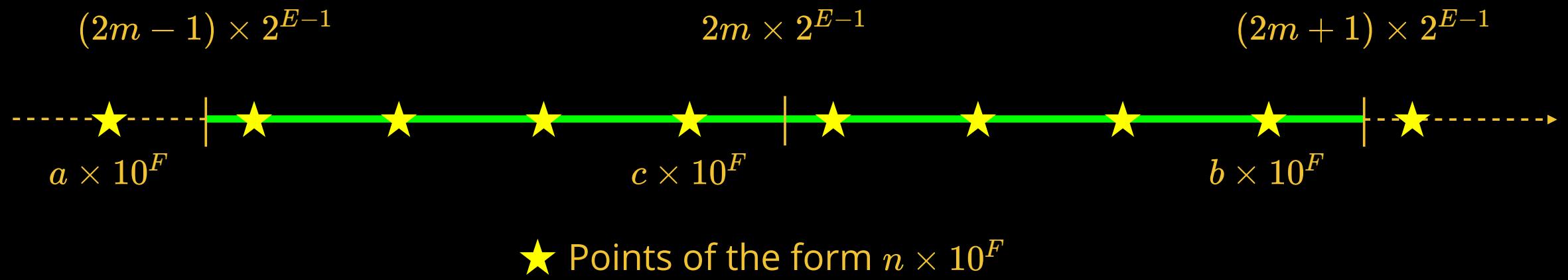


$$a \times 10^F \leq (2m - 1) \times 2^{E-1} < (a + 1) \times 10^F$$

$$a \leq (2m - 1) \times \frac{2^{E-1}}{10^F} < a + 1$$

# Tejú Jaguá - II

---



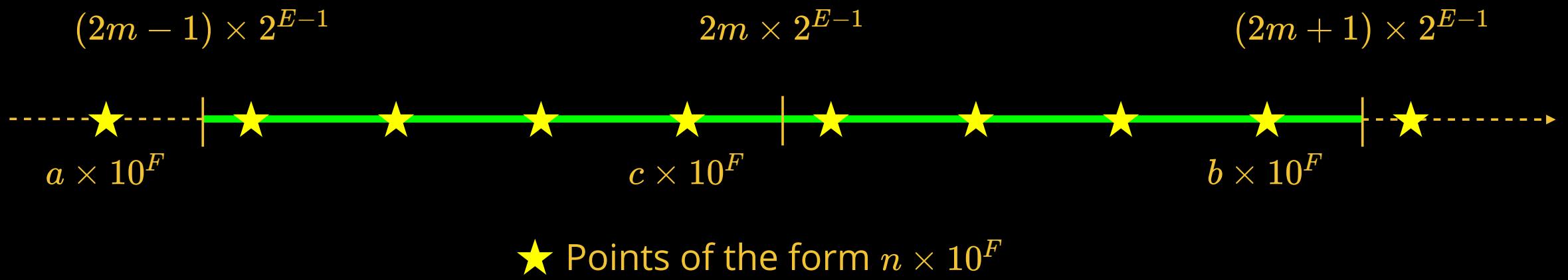
$$a \times 10^F \leq (2m - 1) \times 2^{E-1} < (a + 1) \times 10^F$$

$$a \leq (2m - 1) \times \frac{2^{E-1}}{10^F} < a + 1$$

$$a = \left\lfloor (2m - 1) \times \frac{2^{E-1}}{10^F} \right\rfloor$$

# Tejú Jaguá - II

---



$$a \times 10^F \leq (2m - 1) \times 2^{E-1} < (a + 1) \times 10^F$$

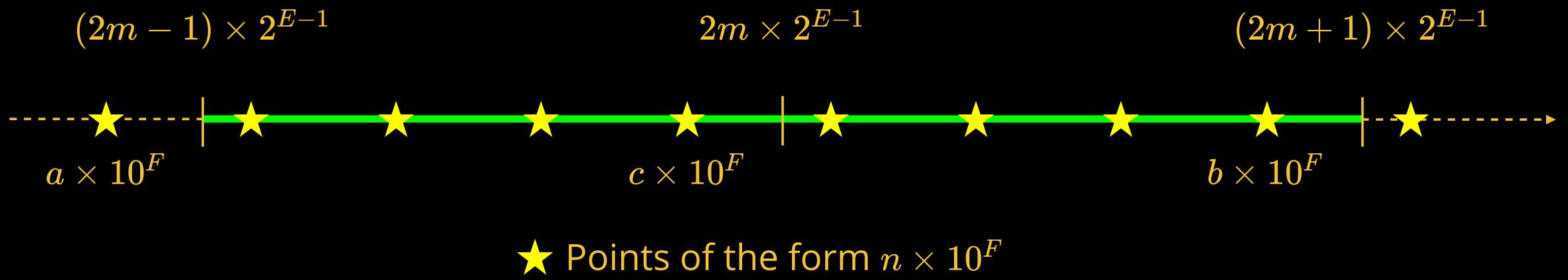
$$a \leq (2m - 1) \times \frac{2^{E-1}}{10^F} < a + 1$$

$$a = \left\lfloor (2m - 1) \times \frac{2^{E-1}}{10^F} \right\rfloor$$

$$b = \left\lfloor (2m + 1) \times \frac{2^{E-1}}{10^F} \right\rfloor$$

# Tejú Jaguá - II

---



$$a \times 10^F \leq (2m - 1) \times 2^{E-1} < (a + 1) \times 10^F$$

$$a \leq (2m - 1) \times \frac{2^{E-1}}{10^F} < a + 1$$

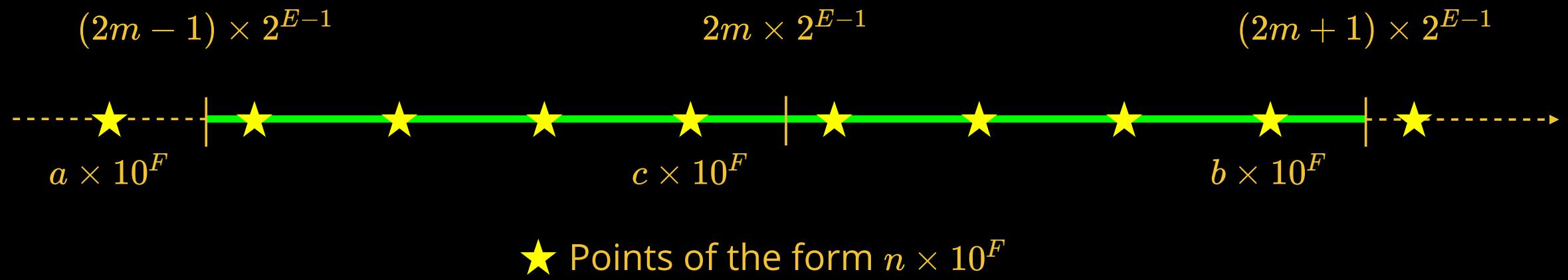
$$a = \left\lfloor (2m - 1) \times \frac{2^{E-1}}{10^F} \right\rfloor$$

$$b = \left\lceil (2m + 1) \times \frac{2^{E-1}}{10^F} \right\rceil$$

$$c = \left\lfloor 2m \times \frac{2^{E-1}}{10^F} \right\rfloor$$

# Tejú Jaguá - II

---



$$a \times 10^F \leq (2m - 1) \times 2^{E-1} < (a + 1) \times 10^F$$

$$a \leq (2m - 1) \times \frac{2^{E-1}}{10^F} < a + 1$$

$$a = \left\lfloor (2m - 1) \times \frac{2^{E-1}}{10^F} \right\rfloor$$

$$b = \left\lfloor (2m + 1) \times \frac{2^{E-1}}{10^F} \right\rfloor$$

$$c = \left\lfloor 2m \times \frac{2^{E-1}}{10^F} \right\rfloor$$

# Tejú Jaguá - III

---

$$(2m - 1) \times 2^{E-1}$$

$$2m \times 2^{E-1}$$

$$(2m + 1) \times 2^{E-1}$$



★ Points of the form  $n \times 10^F$

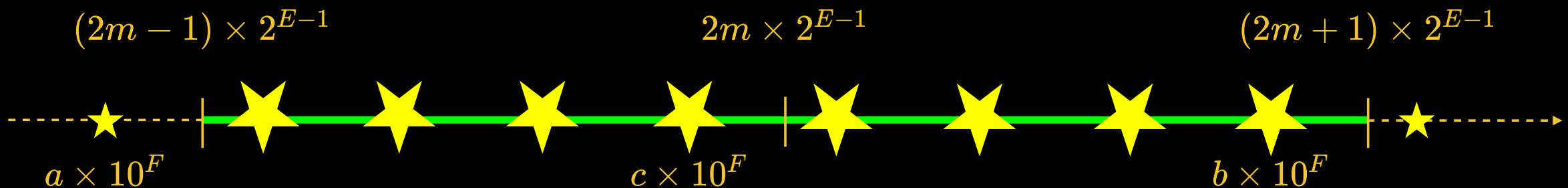
Is any of those ★ of the form  $s \times 10^F$  where  $s$  is a multiple of 10?



If so, then we can do better.

# Tejú Jaguá - III

---



★ Points of the form  $n \times 10^F$

Is any of those ★ of the form  $s \times 10^F$  where  $s$  is a multiple of 10?



If so, then we can do better.

A good **candidate** is  $s = 10 \left\lfloor \frac{b}{10} \right\rfloor$ .

## Return

- the shortest<sup>(\*)</sup>,  $s \times 10^F$ , if you **can**; or
- the closest,  $c \times 10^F$  or  $(c + 1) \times 10^F$ , if you **must**.

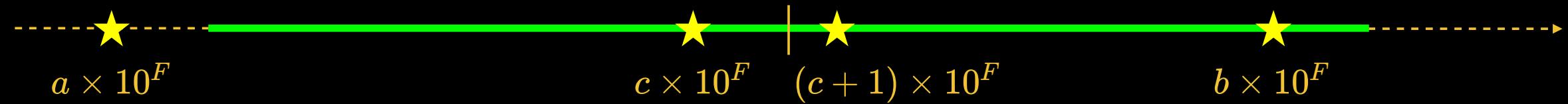
(\*) After removing trailing zeros and adjusting exponent.

# Tejú Jaguá - V

---

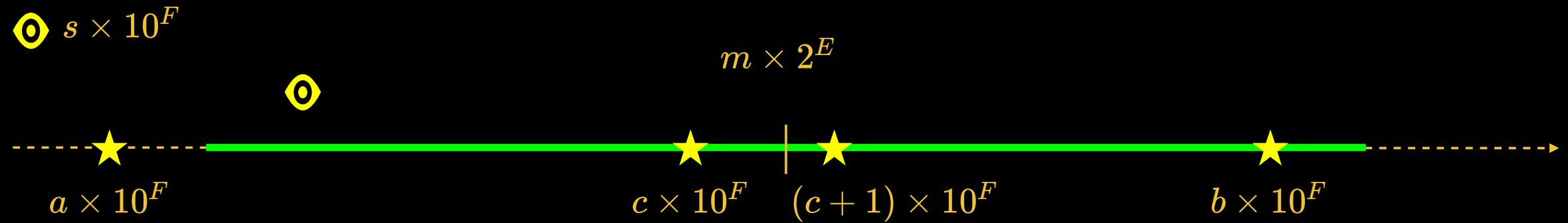
⦿  $s \times 10^F$

$m \times 2^E$



# Tejú Jaguá - V

---



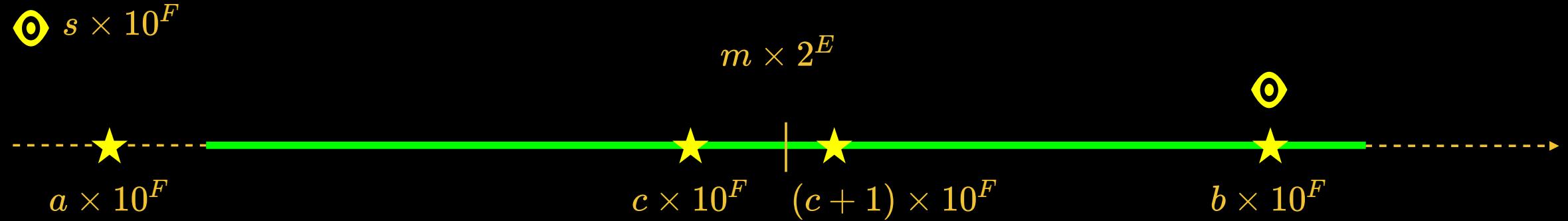
if  $a < s$

if  $s < b$

return the shortest,  $s \times 10^F$

# Tejú Jaguá - V

---



if  $a < s$

if  $s < b$

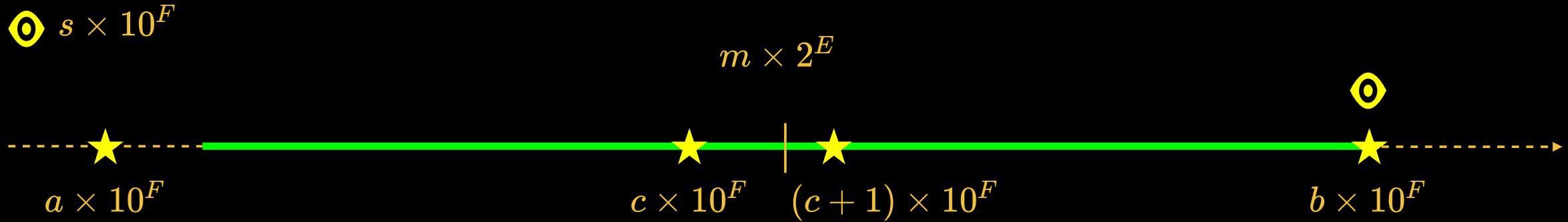
return the shortest,  $s \times 10^F$

if  $b$  is not a tie

return the shortest,  $s \times 10^F$

# Tejú Jaguá - V

---



if  $a < s$

if  $s < b$

return the shortest,  $s \times 10^F$

if  $b$  is not a tie

return the shortest,  $s \times 10^F$

if  $m$  wins against  $m + 1$

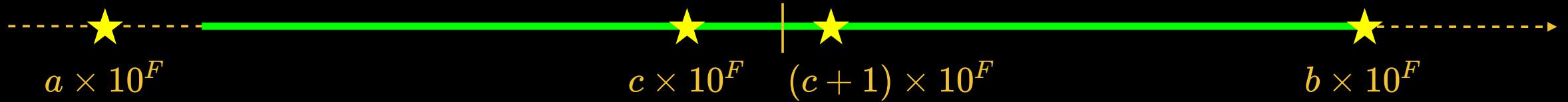
return the shortest,  $s \times 10^F$

# Tejú Jaguá - V

---

①  $s \times 10^F$

$m \times 2^E$



if  $a < s$

if  $s < b$

return the shortest,  $s \times 10^F$

if  $b$  is not a tie

return the shortest,  $s \times 10^F$

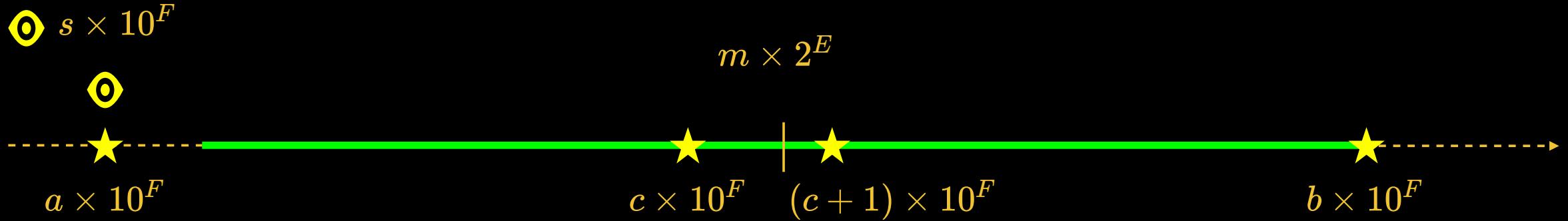
if  $m$  wins against  $m + 1$

return the shortest,  $s \times 10^F$

return the closest,  $c \times 10^F$  or  $(c + 1) \times 10^F$

# Tejú Jaguá - V

---



if  $a < s$

if  $s < b$

return the shortest,  $s \times 10^F$

if  $b$  is not a tie

return the shortest,  $s \times 10^F$

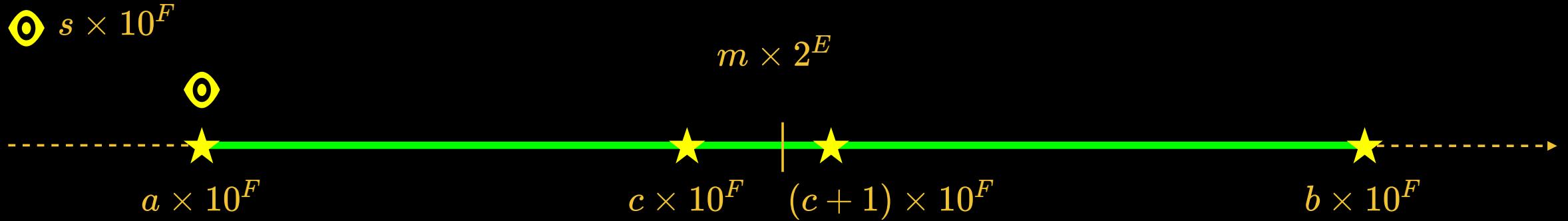
if  $m$  wins against  $m + 1$

return the shortest,  $s \times 10^F$

return the closest,  $c \times 10^F$  or  $(c + 1) \times 10^F$

# Tejú Jaguá - V

---



if  $a < s$

if  $s < b$

return the shortest,  $s \times 10^F$

if  $b$  is not a tie

return the shortest,  $s \times 10^F$

if  $m$  wins against  $m + 1$

return the shortest,  $s \times 10^F$

return the closest,  $c \times 10^F$  or  $(c + 1) \times 10^F$

if  $s = a$

if  $a$  is a tie and  $m$  wins against  $m - 1$

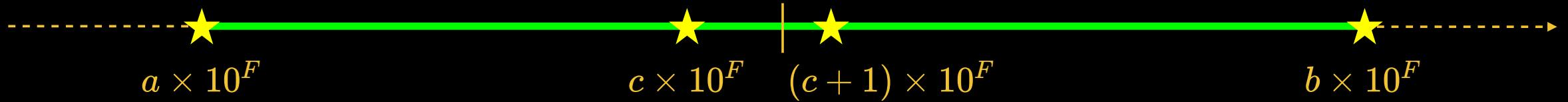
return the shortest,  $s \times 10^F$

# Tejú Jaguá - V

---

①  $s \times 10^F$

$m \times 2^E$



if  $a < s$

if  $s < b$

return the shortest,  $s \times 10^F$

if  $b$  is not a tie

return the shortest,  $s \times 10^F$

if  $m$  wins against  $m + 1$

return the shortest,  $s \times 10^F$

return the closest,  $c \times 10^F$  or  $(c + 1) \times 10^F$

if  $s = a$

if  $a$  is a tie and  $m$  wins against  $m - 1$

return the shortest,  $s \times 10^F$

return the closest,  $c \times 10^F$  or  $(c + 1) \times 10^F$

# Practicalities

---

$$F = \lfloor E \times \log(2) \rfloor = \lfloor E \times 0.301... \rfloor$$

$$F = \left\lfloor \frac{1,292,913,987 \times E}{2^{32}} \right\rfloor$$

Valid  $\forall E \in [-112,815, 112,815]$

# Practicalities

---

$$F = \lfloor E \times \log(2) \rfloor = \lfloor E \times 0.301... \rfloor$$

$$F = \left\lfloor \frac{1,292,913,987 \times E}{2^{32}} \right\rfloor$$

Valid  $\forall E \in [-112,815, 112,815]$

$$a = \left\lfloor M \times \frac{2^{E-1}}{10^F} \right\rfloor$$

$$a = \left\lfloor \frac{M \times M(E)}{2^K} \right\rfloor$$

$K$  = twice the size of floating-point type

$M(E) \cong 2^K \times \frac{2^{E-1}}{10^F}$  can be precomputed  
and stored in a look-up table

# Practicalities

---

$$F = \lfloor E \times \log(2) \rfloor = \lfloor E \times 0.301... \rfloor$$

$$F = \left\lfloor \frac{1,292,913,987 \times E}{2^{32}} \right\rfloor$$

Valid  $\forall E \in [-112,815, 112,815]$

Detecting ties becomes a divisibility check by  $5^F$  using the *minverse* algorithm

$$a = \left\lfloor M \times \frac{2^{E-1}}{10^F} \right\rfloor$$

$$a = \left\lfloor \frac{M \times M(E)}{2^K} \right\rfloor$$

$K$  = twice the size of floating-point type

$M(E) \cong 2^K \times \frac{2^{E-1}}{10^F}$  can be precomputed  
and stored in a look-up table

# Practicalities

---

$$F = \lfloor E \times \log(2) \rfloor = \lfloor E \times 0.301... \rfloor$$

$$F = \left\lfloor \frac{1,292,913,987 \times E}{2^{32}} \right\rfloor$$

Valid  $\forall E \in [-112,815, 112,815]$

Detecting ties becomes a divisibility check by  $5^F$  using the *minverse* algorithm

$$a = \left\lfloor M \times \frac{2^{E-1}}{10^F} \right\rfloor$$

$$a = \left\lfloor \frac{M \times M(E)}{2^K} \right\rfloor$$

$K$  = twice the size of floating-point type

$M(E) \cong 2^K \times \frac{2^{E-1}}{10^F}$  can be precomputed  
and stored in a look-up table

Optimisation for  $c$

# Practicalities

---

$$F = \lfloor E \times \log(2) \rfloor = \lfloor E \times 0.301... \rfloor$$

$$F = \left\lfloor \frac{1,292,913,987 \times E}{2^{32}} \right\rfloor$$

Valid  $\forall E \in [-112,815, 112,815]$

Detecting ties becomes a divisibility check by  $5^F$  using the *minverse* algorithm

Optimisation for integers

$$a = \left\lfloor M \times \frac{2^{E-1}}{10^F} \right\rfloor$$

$$a = \left\lfloor \frac{M \times M(E)}{2^K} \right\rfloor$$

$K$  = twice the size of floating-point type

$M(E) \cong 2^K \times \frac{2^{E-1}}{10^F}$  can be precomputed and stored in a look-up table

Optimisation for  $c$

# Practicalities

---

$$F = \lfloor E \times \log(2) \rfloor = \lfloor E \times 0.301... \rfloor$$

$$F = \left\lfloor \frac{1,292,913,987 \times E}{2^{32}} \right\rfloor$$

Valid  $\forall E \in [-112,815, 112,815]$

Detecting ties becomes a divisibility check by  $5^F$  using the *minverse* algorithm

Optimisation for integers

$$a = \left\lfloor M \times \frac{2^{E-1}}{10^F} \right\rfloor$$

$$a = \left\lfloor \frac{M \times M(E)}{2^K} \right\rfloor$$

$K$  = twice the size of floating-point type

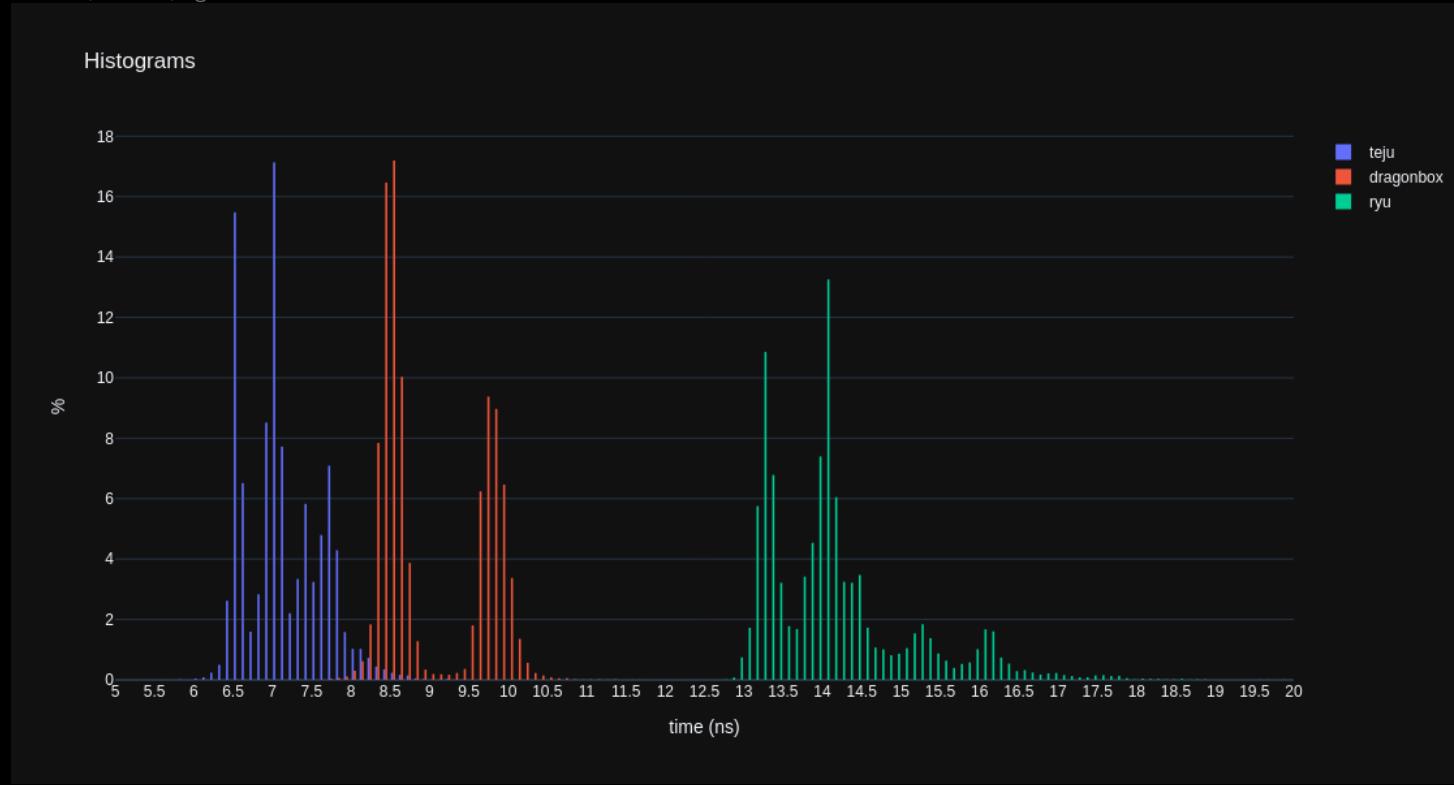
$M(E) \cong 2^K \times \frac{2^{E-1}}{10^F}$  can be precomputed and stored in a look-up table

Optimisation for  $c$

Other tricks

# Benchmarks (centred)

Intel i7 (10510U) - gcc 13.2.1



## teju x dragonbox

wins	99.5%
ties	0.0%
losses	0.5%

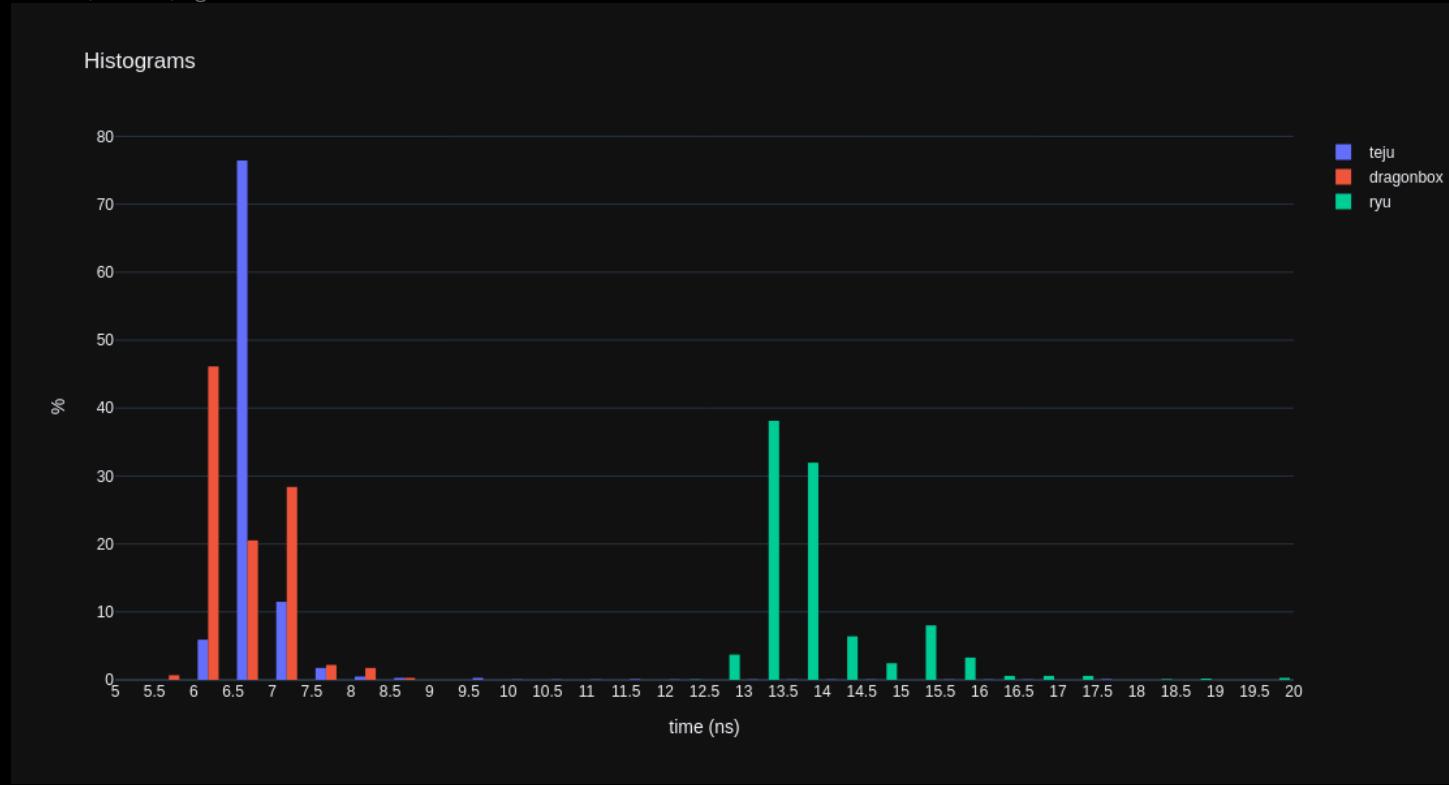
## teju x ryu

wins	100.0%
ties	0.0%
losses	0.0%

	mean	$\sigma$	min	median	max	relative
teju	7.14	0.50	2.84	7.06	9.63	1.00
dragonbox	9.04	0.66	6.52	8.64	11.39	1.27
ryu	14.16	0.98	3.40	14.01	20.50	1.98

# Benchmarks (uncentred)

Intel i7 (10510U) - gcc 13.2.1



## teju x dragonbox

wins	38.2%
ties	0.0%
losses	61.8%

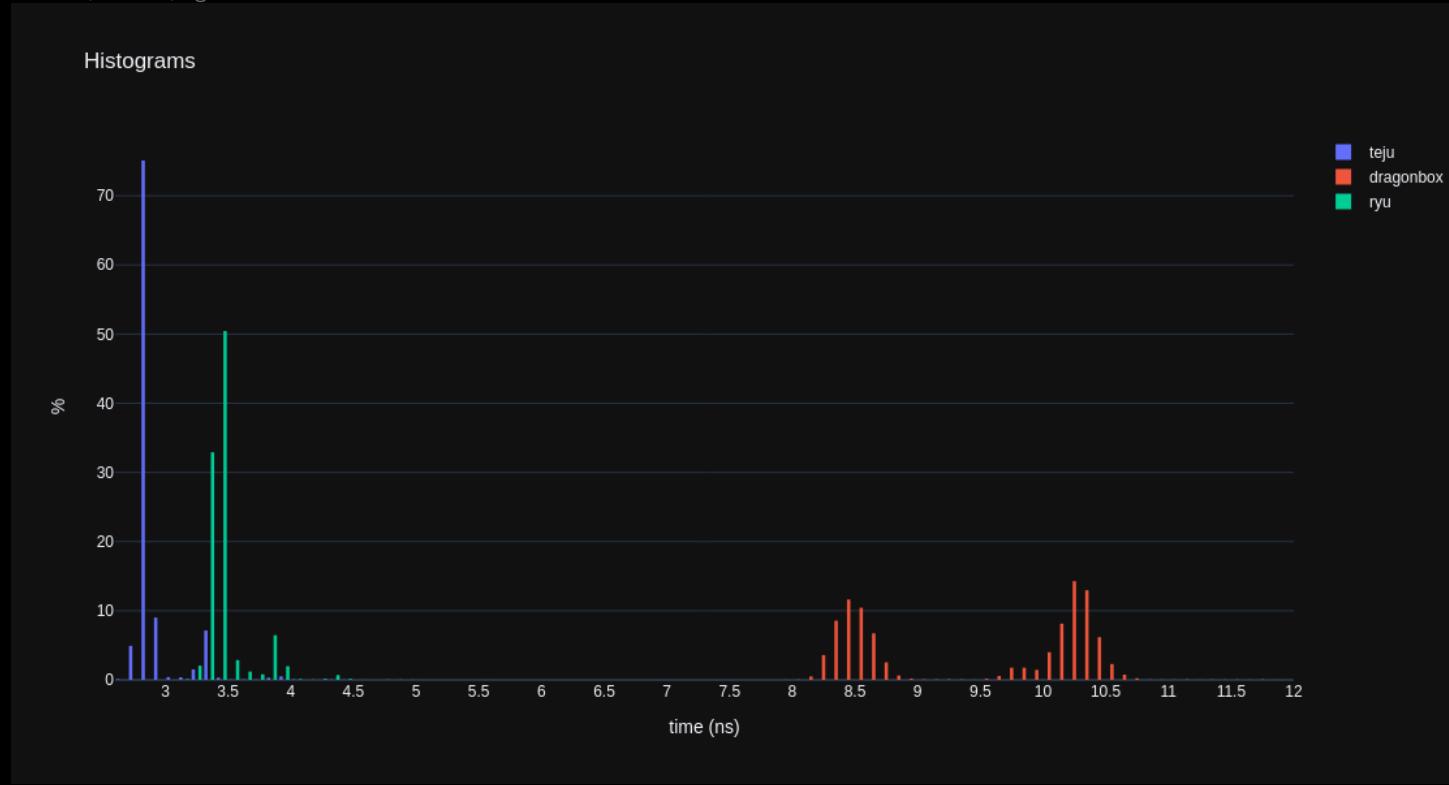
## teju x ryu

wins	100.0%
ties	0.0%
losses	0.0%

	mean	$\sigma$	min	median	max	relative
teju	6.78	0.95	2.79	6.79	17.69	1.00
dragonbox	6.72	0.47	5.95	6.53	8.69	0.99
ryu	13.76	2.79	3.34	13.58	45.12	2.03

# Benchmarks (integers)

Intel i7 (10510U) - gcc 13.2.1



## teju x dragonbox

wins	100.0%
ties	0.0%
losses	0.0%

## teju x ryu

wins	100.0%
ties	0.0%
losses	0.0%

	mean	$\sigma$	min	median	max	relative
teju	2.92	0.18	2.61	2.86	5.05	1.00
dragonbox	9.46	0.90	7.92	10.00	12.58	3.25
ryu	3.47	0.18	3.13	3.42	5.51	1.19

# Joseph Campbell

---



*“The very cave you are afraid to enter turns out to be the source of what you are looking for.”*

# Tejú Jaguá

---



# Tejú Jaguá

---



# Thank you

Cristina Acosta

Lorenz Schneider

Manuel Caicoya

Victor Bogado

