

applied math.
Math 420,

- (a) the model – its construction usually involves simplifications and omissions;
- (b) the data – there may be errors in measuring or estimating values;
- (c) the numerical method – generally based on some approximation;
- (d) the representation of numbers – for example, π cannot be represented exactly by a finite number of digits;
- (e) the arithmetic – frequently errors are introduced in carrying out operations such as addition (+) and multiplication (\times).

numerical
methods.

need to know what errors are in the inputs, but we don't focus on designing measuring equipment that's more accurate.

Science is about quantifying and bounding the errors in the answer...

- Answer is an approximation of truth.
- Science is bounding the error in the answer.

```
$ julia

Documentation: https://docs.julialang.org
Type "?" for help, "]?" for Pkg help.
Version 1.6.7 (2022-07-19)
Official https://julialang.org/ release
```

```
julia> 1+1
2

julia> 1+1-2
0

julia> typeof(1)
Int64
```

↙ exact arithmetic in Julia.

↙ 64 bit integers...

```
julia> 1/2+1/2
1.0

julia> typeof(1.0)
Float64

julia> typeof(1/2+1/2)
Float64

julia> 1/2+1/2-1.0
0.0
```

floating point arithmetic...

← sometimes it's exact

```
julia> x=1/10
0.1

julia> typeof(x)
Float64

julia> y=x+x+x+x+x+x+x+x+x+x
0.9999999999999999
```

← sometimes there is rounding error...

<https://docs.julialang.org/en/v1/manual/integers-and-floating-point-numbers/>

Manual / Integers and Floating-Point Numbers

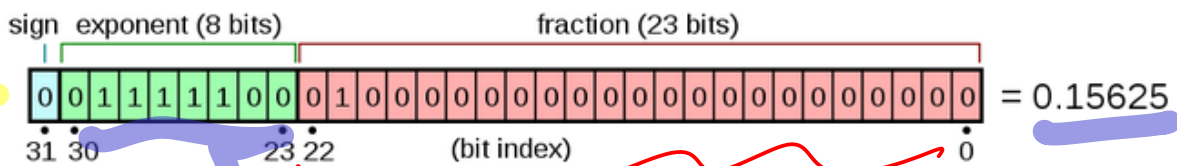
Please read this page...a little...to see the different types of numbers available in Julia.

For Friday read the rest of Step 1...about the pendulum...and pay attention to the types of errors in the story.

The IEEE 754 standard for Floating point...

https://en.wikipedia.org/wiki/IEEE_754

An example of a layout for 32-bit floating point is



and the 64 bit layout is similar.

the first digit is not stored

first non-zero digit in the number.

Scientific notation:

$$0.00476$$

$$4.76 \times 10^{-3}$$

Floating point. Same as scientific notation except in base 2. (only digits are 0 and 1)

$$0.00011011$$

$$1.1011 \times 2^{-4}$$

first non-zero digit is always going to be 1.

$$1.01 \times 2^{e-127}$$

$$= 1.01 \times 2^{124-127}$$

$$= 1.01 \times 2^{-3}$$

exponent in IEEE 754 21
 $e = 1111100 =$

$$4 + 8 + 16 + 32 + 64 = 124$$

$$101 = 5$$

$$1.01 \times 2^{-3} = 101 \times 2^{-5} = \frac{101}{2^5} = 0.15625$$

