Intermezzo Computer representation of numbers

- Mathematically, there are infinitely many real numbers, but computers can only work with a finite subset. Thus, there are gaps between numbers in a computer – this requires rounding. The gaps around the number 1 is **machine epsilon**. Its size depends on the type of number representation. It's about

$$\sim 10^{-16} \text{ for double precission, „double"} \quad (64 \text{ bits})$$
$$\sim 10^{-7} \text{ for single precission „float"} \quad (32 \text{ bits})$$

- Computers store numbers in binary format, i.e., with a base of 2 instead of 10; A "bit" can have value 0, 1 such that

$$1001.101_{(\text{base } 2)} = 8+1+\tfrac{1}{2}+\tfrac{1}{8} = 9.625_{(\text{base } 10)}$$

8, 4, 2, 1, $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$

floating point

- Floating point representation:

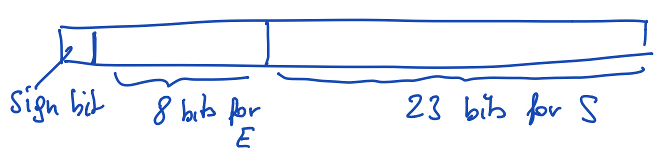
base 10: $x = \pm S \times 10^{E}$ $\quad 1 \le S < 10$

S = *.***…, mantisse; E = exponent

base 2: $x = \pm S \times 2^{E}, \quad 1 \le S < 2$

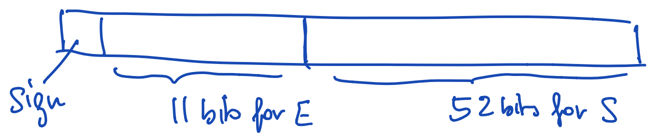
So we only need to store the following:

Single



32 bit

double



64 bit

§ 2.7 Norms and condition numbers

We ask: What consequence has a small perturbation / error in A or b on the solution x of $Ax = b$

Can small changes in A or b potentially have a big influence on x? If yes, when?

Matlab example: $A = \begin{bmatrix} 4.5 & 3.1 \\ 1.6 & 1.1 \end{bmatrix}$ $b = \begin{bmatrix} 19.249 \\ 6.843 \end{bmatrix}$ $b_1 = \begin{bmatrix} 19.25 \\ 6.84 \end{bmatrix}$

$\longrightarrow x = \begin{bmatrix} 3.94 \\ 0.49 \end{bmatrix}$ $x_1 = \begin{bmatrix} 2.9 \\ 2.0 \end{bmatrix}$

Thus: The small change $\Delta b = \begin{bmatrix} 0.001 \\ -0.003 \end{bmatrix}$ has a big influence on x!

We consider the perturbed system

$$\boxed{A(x+\Delta x) = b + \Delta b},$$

where the perturbation Δx is a consequence of the rhs perturbation Δb.

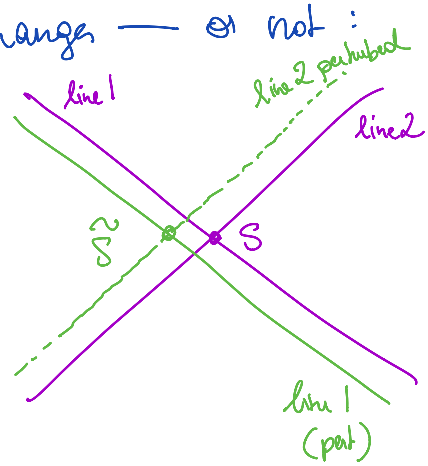
We would like estimates like:

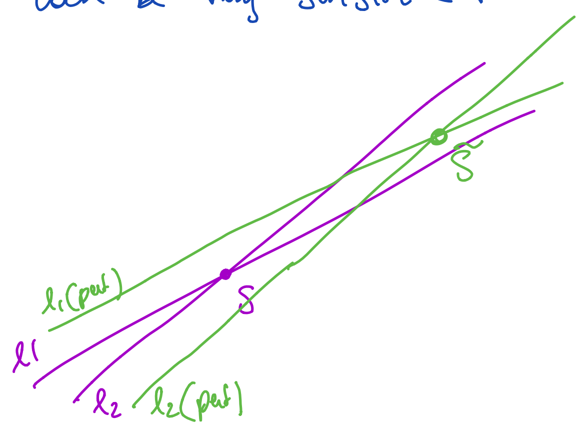
$$\|\Delta x\| \leq K \|\Delta b\| \quad \text{for all } \Delta b \text{ with } K > 0$$

or $$\frac{\|\Delta x\|}{\|x\|} \leq \tilde{\kappa} \frac{\|\Delta b\|}{\|b\|}$$ with $\tilde{\kappa} > 0$

Natural questions:
- How can we find κ, $\tilde{\kappa}$ and how do they depend on A?
- What norms should we use?

Solving a 2x2 linear system is computing the intersection of two lines. This computation can be very sensitive to changes — or not:





⟹ perturbation has a small effect on S

"well-conditioned" systems

⟹ ~~large~~ large influence of perturbations on S

"ill-conditioned" systems

OUR TARGET:

$$\boxed{\kappa \simeq \|A\| \|A^{-1}\|}$$ condition number of A