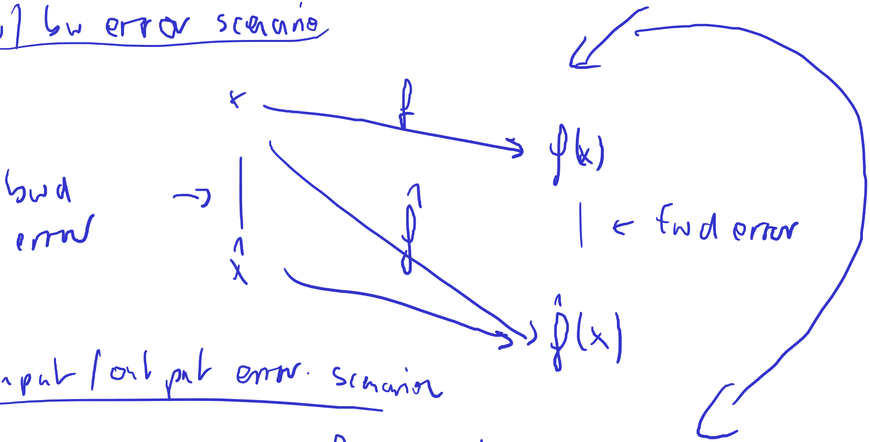
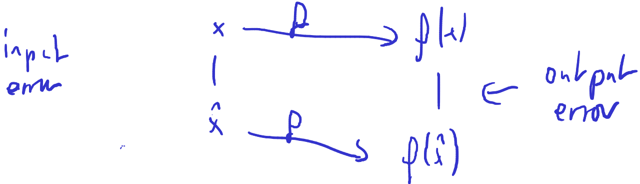
fwd/bw error scenario



input/output error scenario



Sensitivity and Conditioning

Consider a more general setting: An input x and its perturbation $\hat{x}$.

rel. out error $\leq$ κ_{rel} $\cdot$ rel. input error

$$\frac{|f(x) - f(\hat{x})|}{|f(x)|} \leq \kappa_{rel} \cdot \frac{|x - \hat{x}|}{|x|}$$

$$\max \frac{|f(x) - f(\hat{x})|}{|f(x)|} \Big/ \frac{|x - \hat{x}|}{|x|} = \kappa_{rel}$$

For which $x, \hat{x}$? (depends! all scenarios can be relevant)

$(x, \hat{x}) \in S$

Absolute Condition Number

Can you also define an *absolute* condition number?

Absolute Condition Number

rel. error does not work always: $\frac{|0 - \hat{x}|}{0}$

Can you also define an *absolute* condition number?

Certainly:

$$\kappa_{\text{abs}} = \max_{x,\hat{x}} \frac{|f(x) - f(\hat{x})|}{|x - \hat{x}|}$$

But: less commonly used than relative, because we *typically* care about relative error.

When not specified: Assume condition number means *relative*.

Interpreting a Condition Number

What does it mean for condition numbers to be small/large?

cond nr. 5 ←

cond nr. 5000

Relate the (relative) condition number back to the setting of (relative) backward error.

Interpreting a Condition Number

What does it mean for condition numbers to be small/large?

If the condition number is...

- ...small: the problem *well-conditioned* or insensitive
- ...large: the problem *ill-conditioned* or sensitive

Can also talk about condition number for a single input x.

Relate the (relative) condition number back to the setting of (relative) backward error.

rel. output error $\le \kappa \cdot$ rel. input error

Example: Condition Number of Evaluating a Function

$y = f(x)$. Assume f differentiable.

$\Delta x \; (\; x, \hat{x} \qquad \Delta y \; (\; y = f(x), \; \hat{y} = f(\hat{x})$

$$\kappa = \max_{x,\Delta x} \frac{|\Delta y| / |y|}{|\Delta x| / |x|}$$

$$\Delta y = f(x+\Delta x) - f(x) = f'(\xi)\Delta x \approx f'(x)\Delta x$$

$$\kappa \geq \frac{|\Delta y| / |y|}{|\Delta x| / |x|} \approx \frac{|f'(x)|\,|\Delta x| / |f(x)|}{|\Delta x| / |x|} = \frac{|x \cdot f'(x)|}{|f(x)|}$$

Demo: Conditioning of Evaluating tan [cleared]

Stability and Accuracy

Previously: Considered *problems* or *questions*.

Next: Considered *methods*, i.e. computational approaches to find solutions.

When is a method *accurate*?

When is a method *stable*?

Stability and Accuracy

Previously: Considered *problems* or *questions*.

Next: Considered *methods*, i.e. computational approaches to find solutions.

When is a method *accurate*?

> Closeness of method output to true answer for unperturbed input.

When is a method *stable*?

so that we can use conditioning

> - "A method is stable if the result it produces is the exact answer for a nearby input."
> - The above is commonly called *backward stability* and is a stricter requirement than just the temptingly simple:
>
> If the method's sensitivity to variation in the input is no (or not much) greater than that of the problem itself.

Getting into Trouble with Accuracy and Stability

How can I produce inaccurate results?

Getting into Trouble with Accuracy and Stability

How can I produce inaccurate results?

- Apply an inaccurate method
- Apply an unstable method to a well-conditioned problem
- Apply any type of method to an ill-conditioned problem

In-Class Activity: Forward/Backward Error

In-class activity: Forward/Backward Error

$\sin x$

$$\frac{x \cdot f'(x)}{f(x)} = \frac{x \cdot \cos x}{\sin x}$$

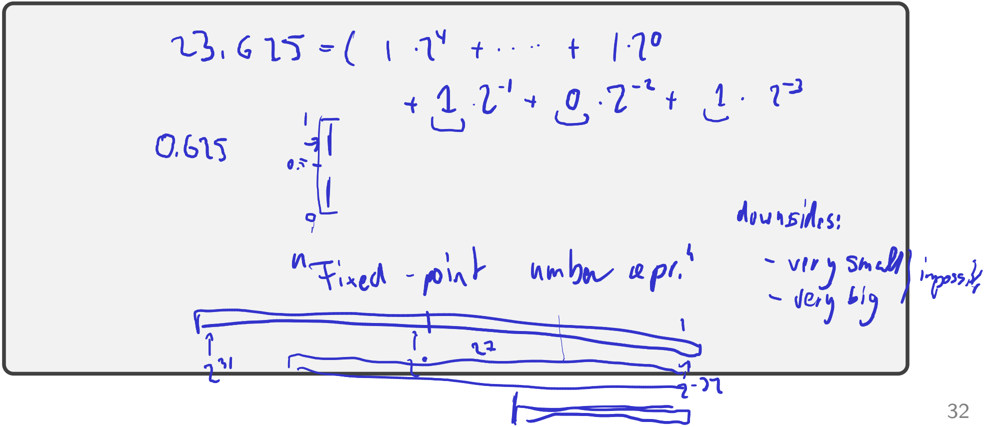
$[0, \frac{\pi}{2})$

Wanted: Real Numbers... in a computer

Computers can represent *integers*, using bits:

$$23 = 1 \cdot 2^4 + 0 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = (10111)_2$$

How would we represent fractions?

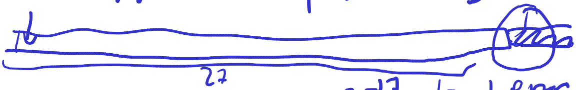


another downside of fixed point:

amount of relative rounding error

depends on magnitude

For. ex. a number that occupies 27 bits



approx rounding error: $2^{-27} = \frac{|\text{error}|}{|\text{number}|}$ error

$23.625 = \underbrace{\qquad\qquad}_{1 \leq \quad < 2} \cdot 2^4$

Floating Point Numbers

Convert $13 = (1101)_2$ into floating point representation.

What pieces do you need to store an FP number?