Mathematical Concepts (G6012)

Lecture 23

Thomas Nowotny
Chichester I, Room CI-105
Office hours: Tuesdays 15:00-16:45
T.Nowotny@sussex.ac.uk

Housekeeping

- MEQ still open – please tell me what what you think!
- Coursework:
  - Turing machine: There are two labels "q_6" by mistake. One of them should be "q_8".
  - "Potato problem": "How often" should have been "How many times" to be precise.

(Thanks everyone who contributed)

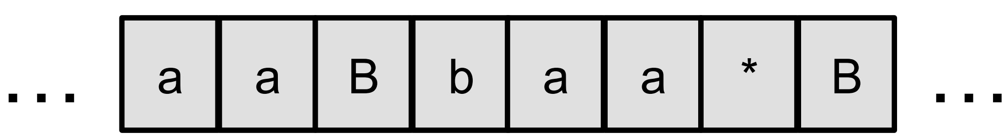
REVISIONS II

TURING MACHINES

Turing Machines (TM)

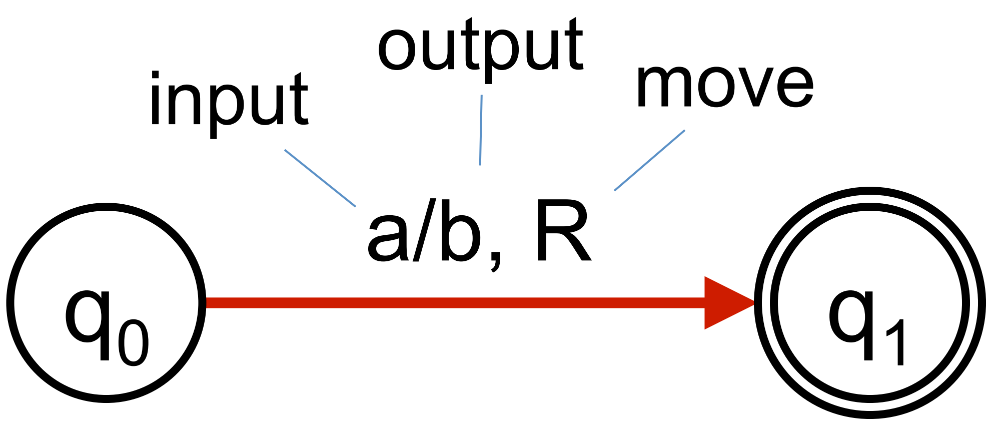
- Are also an extension to FSA
- The main change is to allow **editing the input tape**
- No limit on the size of the tape
- Tape 2-way infinite

- (We will use the symbol B for blank positions)

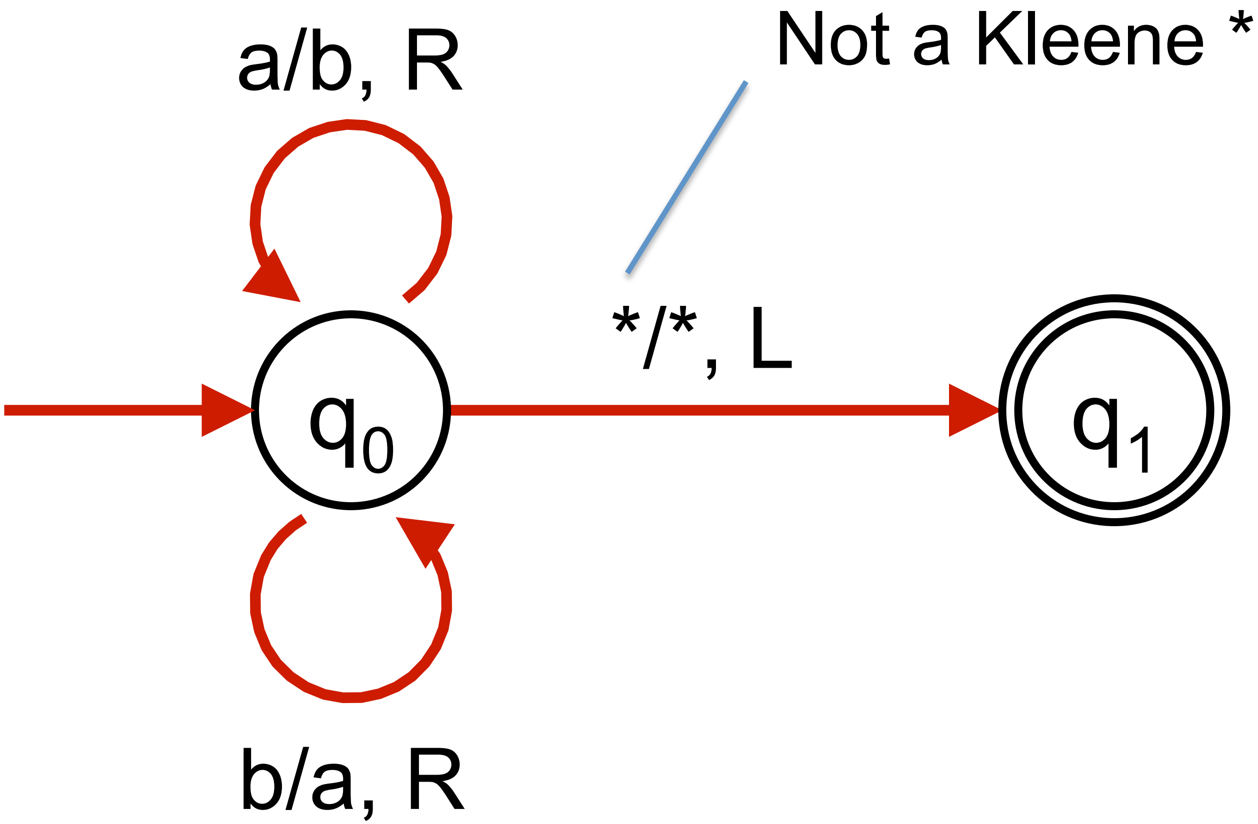


Transitions in TM

- Current state
- New state
- Symbol currently read
- New symbol to replace the read symbol
- Direction to move the tape head (left (L), right (R), stay (S))



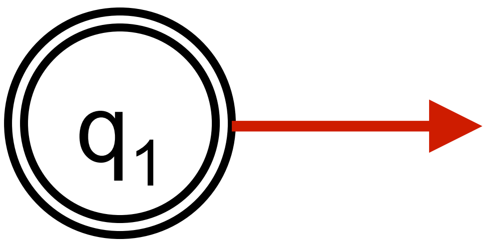
Example



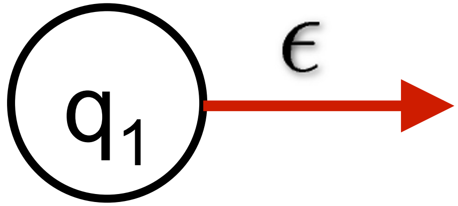
Notes

- TM end in a final state no matter what the state of the tape is
- TM do not consume input
- In TM, there is no role for the empty string
- If anything, Blank strings are written/ read
- TM can be made non-deterministic but doesn’t add power

Common mistakes

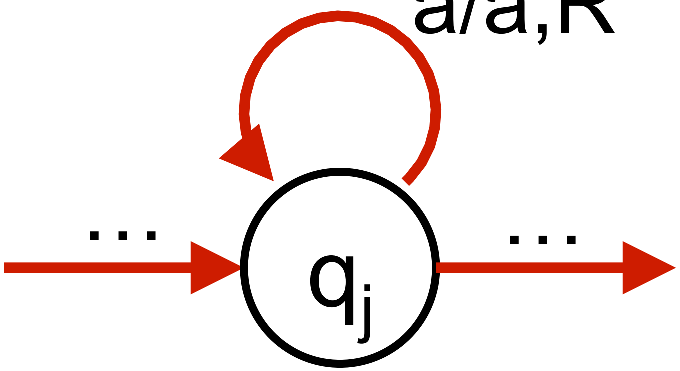


(Rules exiting a final state)

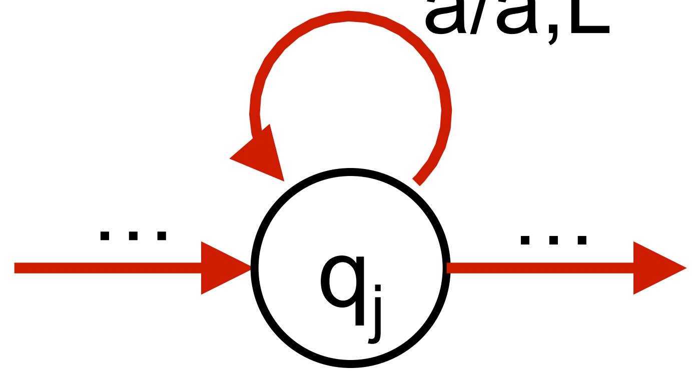


(Rules involving the empty string)

Typical constructs

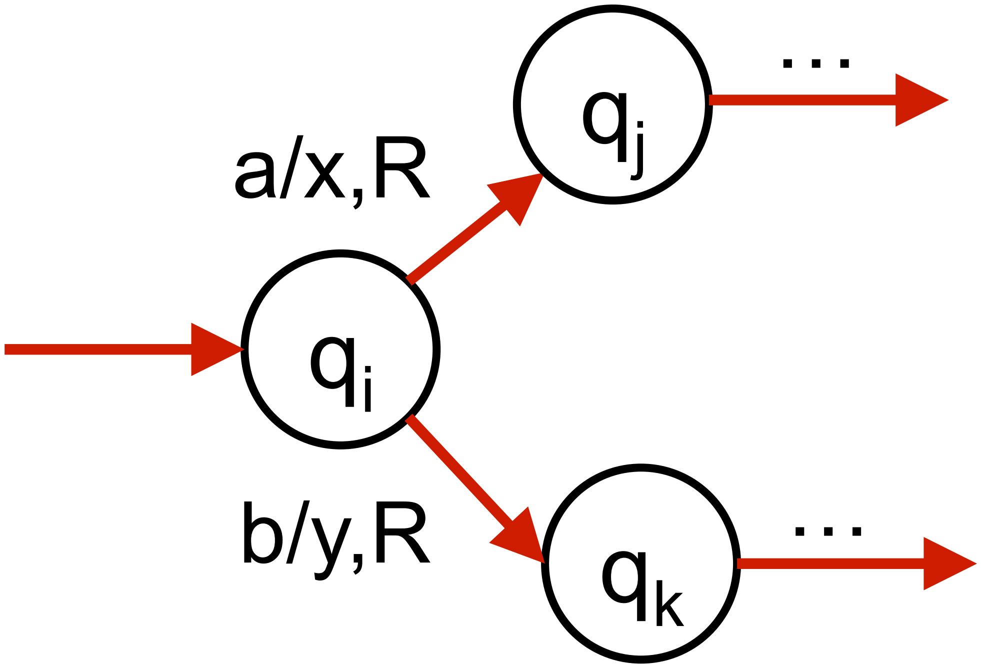


Move tape head
right across all ‘a’



Move tape head
left across all ‘a’

Typical constructs



- Remember an input symbol (through the state)
- Remember what was there (by replacing with something specific)

PROTOCOLS OF COMPUTATION

Any Automation: Protocol of Computation

- Can give "protocol of computation" by making a list of

<table>
<tr><th>FSA</th><th>FST</th><th>PDA</th><th>TM</th></tr>
<tr><td colspan="4">Current state</td></tr>
<tr><td colspan="3">Input tape</td><td>Input/output tape</td></tr>
<tr><td></td><td>Output tape</td><td>Pushdown</td><td></td></tr>
</table>

Example protocols of computation

FSA	FST	PDA	TM
(state, input)	(state, input, output)	(state, input, PD)	(state, input/ output tape)
$(q_0, aabb)$	$(q_0, aabb, \varepsilon)$	$(q_0, aabb, \varepsilon)$	$(q_0, B\underline{a}abbB)$
(q_1, abb)	(q_1, abb, A)	(q_1, abb, A)	$(q_1, Ba\underline{a}bbB)$
(q_1, bb)	(q_1, bb, AA)	(q_1, bb, AA)	$(q_1, Baa\underline{b}bB)$
(q_1, b)	(q_2, b, AAB)	(q_2, b, A)	$(q_2, Baab\underline{b}B)$
(q_2, ε)	$(q_2, \varepsilon, AABB)$	$(q_2, \varepsilon, \varepsilon)$	$(q_2, Baabb\underline{B})$

COMPARISON

Defining automata

- FSA:
  - States q_i
  - Transitions a
  - Input consumed
  - No output
- FST
  - States q_i
  - Transitions a/b
  - Input consumed
  - Output tape
- PDA
  - States qi
  - Transitions a,c/d
  - Input consumed
  - No output
- TM
  - States qi
  - Transitions a/b,R
  - Input just read
  - Output onto same tape

Properties

- FSA:
  - Parse regular languages
  - “least powerful”
- FST
  - Translate regular languages
  - “least powerful”
- PDA
  - Parse context-free languages
  - “medium powerful”
- TM
  - Universal computer: Parsing, translation, etc.
  - “most powerful”

VECTORS AND MATRICES

Vectors: 1D collection of numbers

$$\begin{pmatrix} -1 \\ 3 \\ 9 \end{pmatrix} \in \mathbb{R}^3$$

column vector

arrow, underline

$$\vec{x} = \mathbf{x} = \underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$$

bold

"component" $x_i \in \mathbb{R}$ index

$$\begin{pmatrix} -5 & 3 & 1.1 \end{pmatrix} \in \mathbb{R}^3$$

$$\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \in \mathbb{R}^3$$

Row vector

Similar for all $\mathbb{R}^n$

Matrices: 2D collection of numbers

$$\begin{pmatrix} -1 & 5 & -4 \\ 9.1 & 3 & -4.5 \\ 7 & 0.1 & \sqrt{2} \end{pmatrix} \in \mathcal{M}(3,3)$$ a 3x3 matrix.

capital letter

“entry”, ”element”

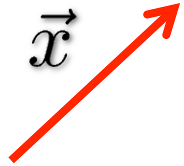
$$A = \underline{\underline{A}} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = (a_{ij}) \in \mathcal{M}(3,3)$$

double underscore

index: row first, column second

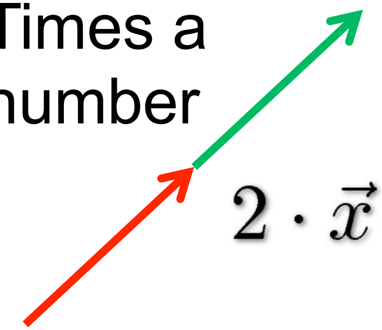
Vectors: Operations and geometric interpretation

Vector $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$



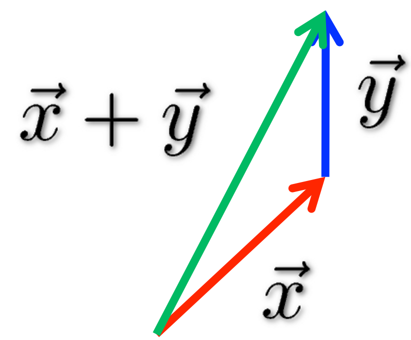
Arrow in space

Times a number



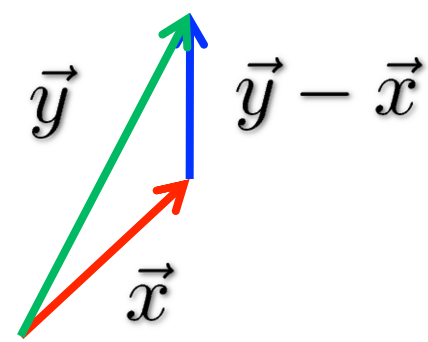
Sum of vectors

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix}$$



Difference of vectors

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 - x_1 \\ y_2 - x_2 \end{pmatrix}$$



Matrices: Operations

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

$$k \cdot A = \begin{pmatrix} ka_{11} & ka_{12} \\ ka_{21} & ka_{22} \end{pmatrix}$$

$$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix}$$

$$A - B = \begin{pmatrix} a_{11} - b_{11} & a_{12} - b_{12} \\ a_{21} - b_{21} & a_{22} - b_{22} \end{pmatrix}$$

Matrices: Operations

$$A \cdot B = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

Written in components:

$$(k \cdot A)_{ij} = k a_{ij}$$

$$(A + B)_{ij} = a_{ij} + b_{ij}$$

$$(A - B)_{ij} = a_{ij} - b_{ij}$$

$$(A \cdot B)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

We can multiply matrices if the inner dimensions agree.

Properties of matrix operations

Associativity

$$A + B + C = (A + B) + C = A + (B + C)$$

Commutativity $A + B = B + A$

Scalar multiplication $r \cdot (A + B) = r \cdot A + r \cdot B$

$$(r + s) \cdot A = r \cdot A + s \cdot A$$

Multiplication:

Associativity $A \cdot (B \cdot C) = (A \cdot B) \cdot C$

But NOT Commutativity: $A \cdot B \neq B \cdot A$ Except for special cases)

Matrix-vector multiplication

$$A\vec{x} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} =$$

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{pmatrix}$$

The result is a vector.

Properties of Matrix-Vector Multiplication

Linear (both ways)

$$A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y}$$

$$(A + B)\vec{x} = A\vec{x} + B\vec{x}$$

Associative:

$$(A \cdot B)\vec{x} = A(B\vec{x})$$

Interpretation

Matrices are transformations (linear functions)

$$A = \underline{\underline{A}} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \quad \text{M(3,3)}$$

$$A : \mathbb{R}^3 \to \mathbb{R}^3$$

$$\vec{x} \mapsto A \cdot \vec{x}$$

$$A\,\vec{x}$$

A maps vectors from $\mathbb{R}^3$ to vectors in $\mathbb{R}^3$

$A \cdot \vec{x}$, $A\,\vec{x}$: Matrix- vector multiplication

How do we see what a matrix does geometrically?

The columns of the matrix are the images of the basis vectors

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \qquad \vec{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

All vectors can be expressed as a combination of these.
So, knowing where they are mapped to tells us everything about the geometric action of a matrix.

Matrix transpose

Transposition is the operation where lines and columns are swapped. Or a reflection along the diagonal, if you want:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}^T = \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}$$

$(A^T)_{ij} = a_{ji}$ where we assume $(A)_{ij} = a_{ij}$

And we saw $(A \cdot B)^T = B^T A^T$ (Note the order!)

Scalar product

The scalar product of two vectors is defined as

$$\vec{x} \cdot \vec{y} := \sum_{i=1}^{3} x_i y_i$$ also denoted as $\langle \vec{x}, \vec{y} \rangle$

It is a special case of Matrix multiplication:

$$\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = x_1 y_1 + x_2 y_2 + x_3 y_3$$

Has to do with length and angles:

$$\vec{x} \cdot \vec{y} = ||\vec{x}|| \; ||\vec{y}|| \cos \alpha_{x,y}$$

Length and distances

- Euclidean norm (length)

$\vec{x} \in \mathbb{R}^n$

Norm of $\vec{x}$ is $||\vec{x}|| := \sqrt{\sum_{i=1}^{n} x_i^2} = \sqrt{\vec{x} \cdot \vec{x}}$

Multiplication with number

$||k \cdot \vec{x}|| = |k| \cdot ||\vec{x}||$

Triangle inequality

$||\vec{x} + \vec{y}|| \leq ||\vec{x}|| + ||\vec{y}||$

- Euclidean distance

$d(\vec{x}, \vec{y}) = ||\vec{x} - \vec{y}||$

CALCULUS

Functions

domain of the function

co-domain or image of the function

function name

$$f : \mathbb{R} \rightarrow \mathbb{R}_+$$

argument of *f*

$$x \mapsto 5x^2 + 1$$

alternative notation:

$$f(x) = 5x^2 + 1$$

Functions are also often called map, mapping or transformation.

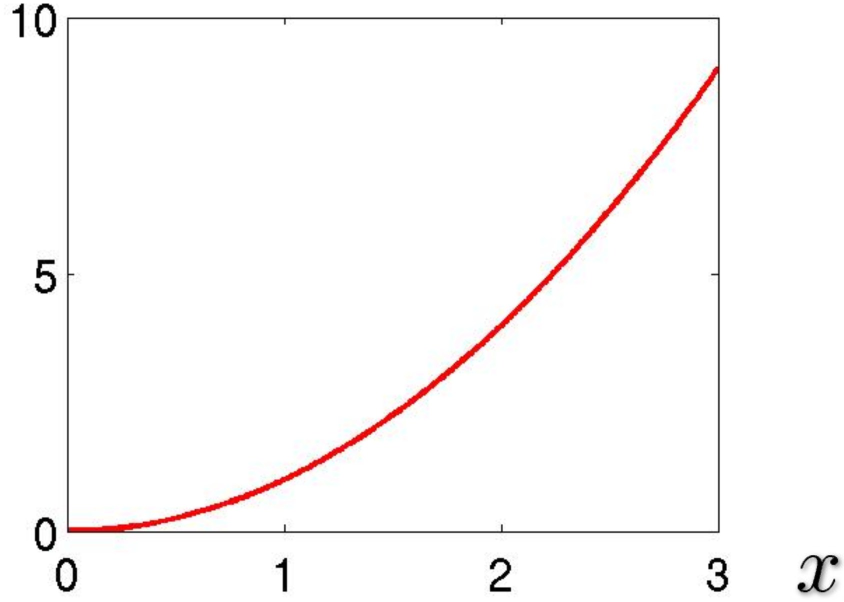
Examples of functions

- Linear functions (straight lines)
- Polynomials (powers of x)
- Trigonometric functions (sin(x), cos(x))
- Exponential functions (exp(x))
- Logarithm (log(x))

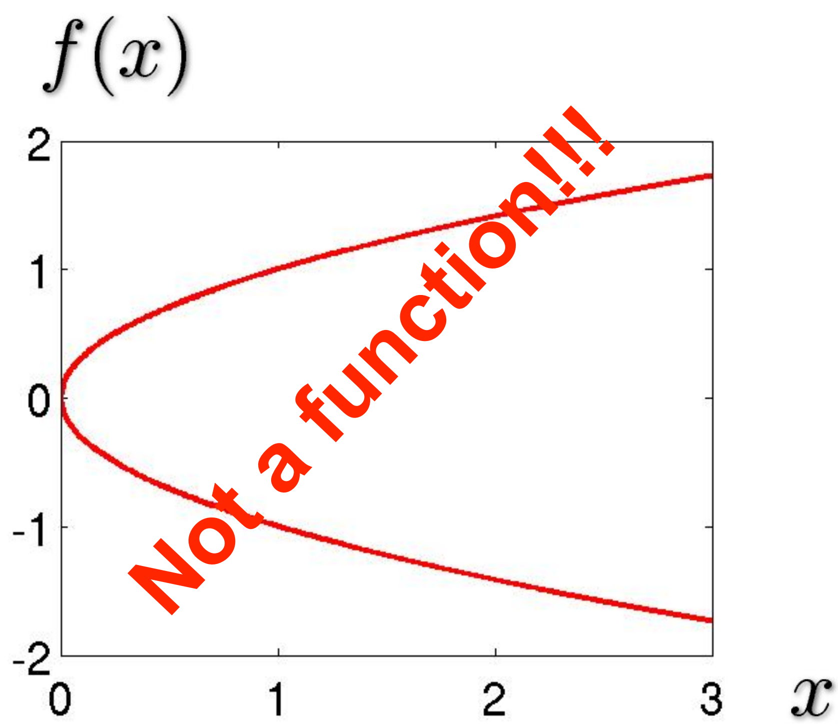
- Matrices (=high-dimensional linear functions)

Graph of a function

$$f(x) = x^2$$



A parabola



Limits

Problem: Sequence of numbers, e.g.

$$a_n = \frac{1}{n} \quad \text{and} \quad n = 1, 2, \ldots$$

What happens to $\frac{1}{n}$ if we make *n* ever larger?

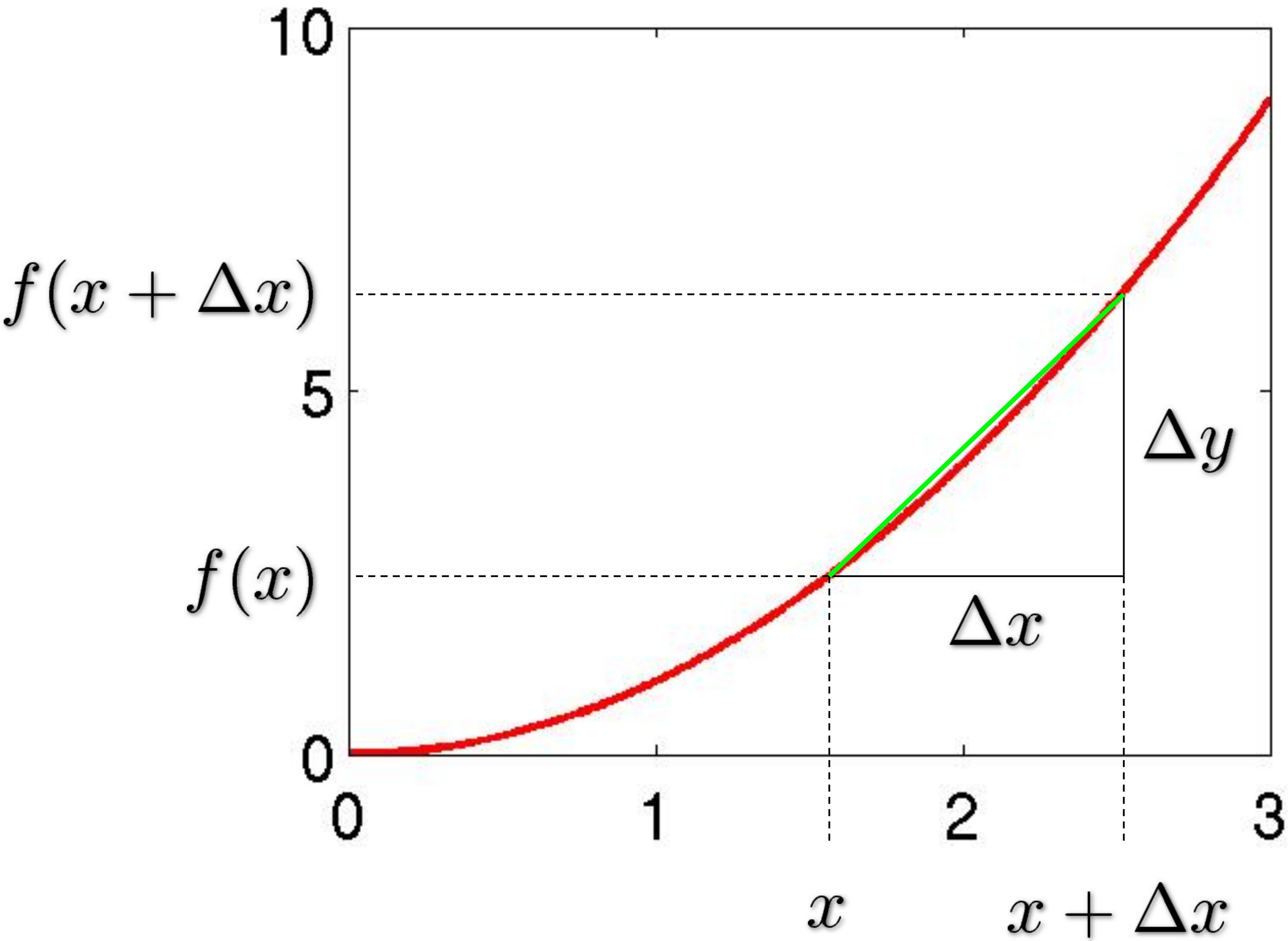
Mathematicians write $\lim_{n \to \infty} \frac{1}{n} = 0$

Two *typical* cases:
The numbers go ever larger (the series diverges), or (more interesting) it converges to a number (here 0). Sometimes it does neither.

Formal Definition

A sequence a_n converges to a number $x \in \mathbb{R}$ if for all (small) $\epsilon > 0$ there is a number $N \in \mathbb{N}$ such that $|a_n - x| < \epsilon$ for all $n \geq N$.

Derivatives: Idea



Derivative of a smooth function

- The derivative of a smooth function is the value the ratio

$$\frac{f(x+\Delta x)-f(x)}{\Delta x}$$

converges to for smaller and smaller

Δx , mathematicians write

$$f'(x)=\lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}$$

Derivatives: Basic rules

Rule name	Function	Derivative
Polynomials	$f(x) = x^n$	$f'(x) = n\,x^{n-1}$
Constant factor	$g(x) = a\,f(x)$	$g'(x) = a\,f'(x)$
Sum and Difference	$h(x) = f(x) + g(x)$	$h'(x) = f'(x) + g'(x)$

Special functions

Function	Derivative
$\exp(x) = e^x$	$\exp(x) = e^x$
$\log(x) = \ln(x)$	$\frac{1}{x}$
$\sin(x)$	$\cos(x)$
$\cos(x)$	$-\sin(x)$

Derivatives: Rules

Product rule

Function $h(x) = f(x) \cdot g(x)$

Derivative $h'(x) = f'(x) \cdot g(x) + f(x) \cdot g'(x)$

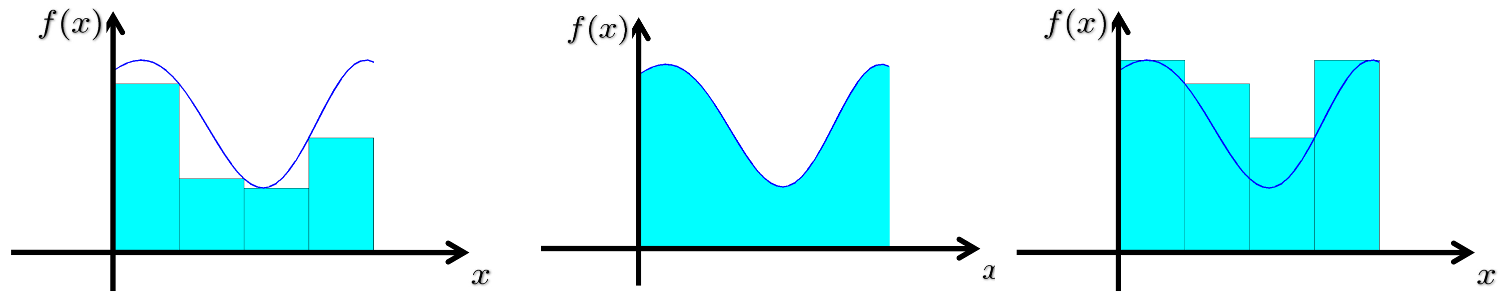
Chain rule

Function $h(x) = f\big(g(x)\big)$ $h = f \circ g$

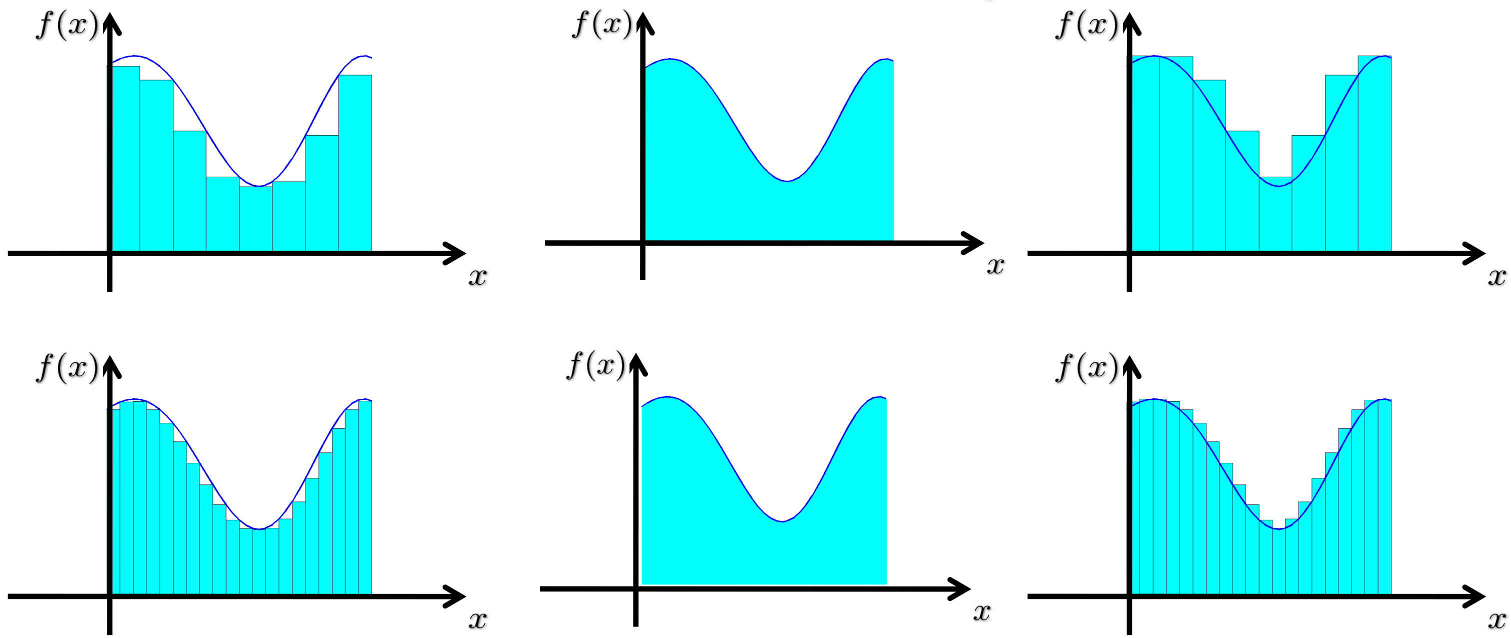
Derivative $h'(x) = f'\big(g(x)\big) \cdot g'(x)$

$$h' = f' \circ g \cdot g'$$

Integration: Idea

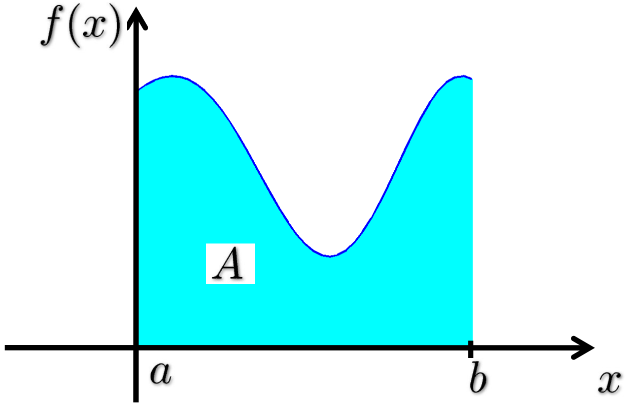


$$A_- \leq A \leq A_+$$



(Riemann) integral

$$A = \lim_{\Delta x \to 0} A_- = \lim_{\Delta x \to 0} A_+$$



is called “(Riemann) integral”

Notation:

$$A = \int_a^b f(x)\, dx$$

upper limit

integration variable
(dummy variable)

Integral sign

lower limit

Main theorem of differential and integral calculus

In principle, one could calculate integrals from first principles, but fortunately…

Integration is the opposite of differentiation!

Special functions

Function	Integral
$f(x) = x^n$	$\int_0^x t^n dt = \frac{1}{n+1} x^{n+1}$
$\exp(x) = e^x$	$\exp(x) = e^x$
$\frac{1}{x}$	$\log(x) = \ln(x)$
$\cos(x)$	$\sin(x)$
$\sin(x)$	$-\cos(x)$

Practical tips

$$\int_a^b f(x)\,dx = \Big[\,F(x)\,\Big]_a^b \; = F(b) - F(a)$$

"anti-derivative"

Example:

$$\int_2^3 x^3\,dx = \;\Big[\;\frac{1}{4}x^4\;\Big]_2^3 = \;\frac{1}{4}3^4 - \frac{1}{4}2^4 =$$

$$\frac{81-16}{4} = \frac{65}{4}$$