Part II

Foundations

3 Goals

- Gain knowledge about efficient algorithms for important problems, i.e., learn how to solve certain types of problems efficiently.
- Learn how to analyze and judge the efficiency of algorithms.
- Learn how to design efficient algorithms.

4 Modelling Issues

What do you measure?

- Memory requirement
- Running time
- Number of comparisons
- Number of multiplications
- Number of hard-disc accesses
- Program size
- Power consumption
- ...

4 Modelling Issues

How do you measure?

- Implementing and testing on representative inputs
  - How do you choose your inputs?
  - May be very time-consuming.
  - Very reliable results if done correctly.
  - Results only hold for a specific machine and for a specific set of inputs.
- Theoretical analysis in a specific model of computation.
  - Gives asymptotic bounds like “this algorithm always runs in time $\mathcal{O}(n^2)$”.
  - Typically focuses on the worst case.
  - Can give lower bounds like “any comparison-based sorting algorithm needs at least $\Omega(n \log n)$ comparisons in the worst case”.

4 Modelling Issues

Input length

The theoretical bounds are usually given by a function $f : \mathbb{N} \to \mathbb{N}$ that maps the input length to the running time (or storage space, comparisons, multiplications, program size etc.).

The input length may e.g. be

- the size of the input (number of bits)
- the number of arguments

Example 1

Suppose n numbers from the interval $\{1, \dots, N\}$ have to be sorted. In this case we usually say that the input length is n instead of e.g. $n \log N$, which would be the number of bits required to encode the input.

Model of Computation

How to measure performance

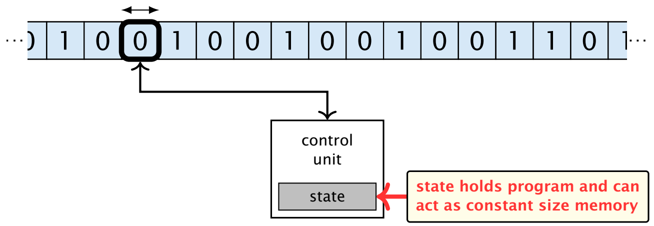
1. Calculate running time and storage space etc. on a simplified, idealized model of computation, e.g. Random Access Machine (RAM), Turing Machine (TM), ...
2. Calculate number of certain basic operations: comparisons, multiplications, harddisc accesses, ...

Version 2. is often easier, but focusing on one type of operation makes it more difficult to obtain meaningful results.

Turing Machine

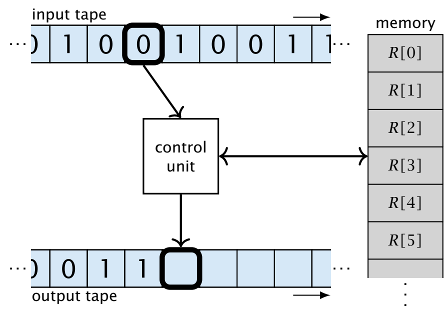
- Very simple model of computation.
- Only the “current” memory location can be altered.
- Very good model for discussing computabiliy, or polynomial vs. exponential time.
- Some simple problems like recognizing whether input is of the form xx, where x is a string, have quadratic lower bound.

$\Rightarrow$ **Not a good model for developing efficient algorithms.**



Random Access Machine (RAM)

- Input tape and output tape (sequences of zeros and ones; unbounded length).
- Memory unit: infinite but countable number of registers $R[0], R[1], R[2], \dots$.
- Registers hold integers.
- Indirect addressing.



Note that in the picture on the right the tapes are one-directional, and that a READ- or WRITE-operation always advances its tape.

Random Access Machine (RAM)

Operations

- input operations (input tape → $R[i]$)
  - READ i
- output operations ($R[i]$ → output tape)
  - WRITE i
- register-register transfers
  - $R[j]$:= $R[i]$
  - $R[j]$:= 4
- **indirect** addressing
  - $R[j]$:= $R[R[i]]$
    loads the content of the $R[i]$-th register into the j-th register
  - $R[R[i]] := R[j]$
    loads the content of the j-th into the $R[i]$-th register

Random Access Machine (RAM)

Operations

- branching (including loops) based on comparisons
  - jump x
    jumps to position x in the program;
    sets instruction counter to x;
    reads the next operation to perform from register $R[x]$
  - jumpz x $R[i]$
    jump to x if $R[i] = 0$
    if not the instruction counter is increased by 1;
  - jumpi i
    jump to $R[i]$ (indirect jump);
- arithmetic instructions: $+$, $-$, $\times$, $/$
  - $R[i]$:= $R[j]$ + $R[k]$;
    $R[i]$:= -$R[k]$;

The jump-directives are very close to the jump-instructions contained in the assembler language of real machines.

Model of Computation

- uniform cost model
  Every operation takes time 1.
- logarithmic cost model
  The cost depends on the content of memory cells:
  - The time for a step is equal to the largest operand involved;
  - The storage space of a register is equal to the length (in bits) of the largest value ever stored in it.

Bounded word RAM model: cost is uniform but the largest value stored in a register may not exceed 2^w, where usually $w = \log_2 n$.

The latter model is quite realistic as the word-size of a standard computer that handles a problem of size n must be at least $\log_2 n$ as otherwise the computer could either not store the problem instance or not address all its memory.

4 Modelling Issues

Example 2

Algorithm 1 RepeatedSquaring(n)

```
1: r ← 2;
2: for i = 1 → n do
3:     r ← r^2
4: return r
```

- running time (for Line 3):
  - uniform model: n steps
  - logarithmic model:
    $2 + 3 + 5 + \cdots + (1 + 2^n) = 2^{n+1} - 1 + n = \Theta(2^n)$
- space requirement:
  - uniform model: $\mathcal{O}(1)$
  - logarithmic model: $\mathcal{O}(2^n)$

There are different types of complexity bounds:

- best-case complexity:

$$C_{bc}(n) := \min\{C(x) \mid |x| = n\}$$

  Usually easy to analyze, but not very meaningful.
- worst-case complexity:

$$C_{wc}(n) := \max\{C(x) \mid |x| = n\}$$

  Usually moderately easy to analyze; sometimes too pessimistic.
- average case complexity:

$$C_{avg}(n) := \frac{1}{|I_n|} \sum_{|x|=n} C(x)$$

  more general: probability measure μ

$$C_{avg}(n) := \sum_{x \in I_n} \mu(x) \cdot C(x)$$

μ is a probability distribution over inputs of length n.

$C(x)$	cost of instance x
$\lvert x\rvert$	input length of instance x
I_n	set of instances of length n

There are different types of complexity bounds:

- amortized complexity:
  The average cost of data structure operations over a worst case sequence of operations.
- randomized complexity:
  The algorithm may use random bits. Expected running time (over all possible choices of random bits) for a fixed input x. Then take the worst-case over all x with $|x| = n$.

μ is a probability distribution over inputs of length n.

$C(x)$	cost of instance x
$\lvert x\rvert$	input length of instance x
I_n	set of instances of length n

4 Modelling Issues

Bibliography

[MS08] Kurt Mehlhorn, Peter Sanders:
Algorithms and Data Structures — The Basic Toolbox,
Springer, 2008

[CLRS90] Thomas H. Cormen, Charles E. Leiserson, Ron L. Rivest, Clifford Stein:
Introduction to algorithms (3rd ed.),
McGraw-Hill, 2009

Chapter 2.1 and 2.2 of [MS08] and Chapter 2 of [CLRS90] are relevant for this section.

5 Asymptotic Notation

We are usually not interested in exact running times, but only in an asymptotic classification of the running time, that ignores constant factors and constant additive offsets.

- We are usually interested in the running times for large values of n. Then constant additive terms do not play an important role.
- An exact analysis (e.g. *exactly* counting the number of operations in a RAM) may be hard, but wouldn't lead to more precise results as the computational model is already quite a distance from reality.
- A linear speed-up (i.e., by a constant factor) is always possible by e.g. implementing the algorithm on a faster machine.
- Running time should be expressed by simple functions.

Asymptotic Notation

Formal Definition

Let f, g denote functions from $\mathbb{N}$ to $\mathbb{R}^+$.

- $\mathcal{O}(f) = \{g \mid \exists c > 0\ \exists n_0 \in \mathbb{N}_0\ \forall n \geq n_0 : [g(n) \leq c \cdot f(n)]\}$
  (set of functions that asymptotically grow not faster than f)
- $\Omega(f) = \{g \mid \exists c > 0\ \exists n_0 \in \mathbb{N}_0\ \forall n \geq n_0 : [g(n) \geq c \cdot f(n)]\}$
  (set of functions that asymptotically grow not slower than f)
- $\Theta(f) = \Omega(f) \cap \mathcal{O}(f)$
  (functions that asymptotically have the same growth as f)
- $o(f) = \{g \mid \forall c > 0\ \exists n_0 \in \mathbb{N}_0\ \forall n \geq n_0 : [g(n) \leq c \cdot f(n)]\}$
  (set of functions that asymptotically grow slower than f)
- $\omega(f) = \{g \mid \forall c > 0\ \exists n_0 \in \mathbb{N}_0\ \forall n \geq n_0 : [g(n) \geq c \cdot f(n)]\}$
  (set of functions that asymptotically grow faster than f)

Asymptotic Notation

There is an equivalent definition using limes notation (assuming that the respective limes exists). f and g are functions from $\mathbb{N}_0$ to $\mathbb{R}_0^+$.

- $g \in \mathcal{O}(f)$: $\quad 0 \leq \lim_{n\to\infty} \frac{g(n)}{f(n)} < \infty$
- $g \in \Omega(f)$: $\quad 0 < \lim_{n\to\infty} \frac{g(n)}{f(n)} \leq \infty$
- $g \in \Theta(f)$: $\quad 0 < \lim_{n\to\infty} \frac{g(n)}{f(n)} < \infty$
- $g \in o(f)$: $\quad \lim_{n\to\infty} \frac{g(n)}{f(n)} = 0$
- $g \in \omega(f)$: $\quad \lim_{n\to\infty} \frac{g(n)}{f(n)} = \infty$

- Note that for the version of the Landau notation defined here, we assume that f and g are positive functions.
- There also exist versions for arbitrary functions, and for the case that the limes is not infinity.

Asymptotic Notation

Abuse of notation

1. People write $f = \mathcal{O}(g)$, when they mean $f \in \mathcal{O}(g)$. This is **not** an equality (how could a function be equal to a set of functions).
2. People write $f(n) = \mathcal{O}(g(n))$, when they mean $f \in \mathcal{O}(g)$, with $f : \mathbb{N} \to \mathbb{R}^+, n \mapsto f(n)$, and $g : \mathbb{N} \to \mathbb{R}^+, n \mapsto g(n)$.
3. People write e.g. $h(n) = f(n) + o(g(n))$ when they mean that there exists a function $z : \mathbb{N} \to \mathbb{R}^+, n \mapsto z(n), z \in o(g)$ such that $h(n) = f(n) + z(n)$.

2. In this context $f(n)$ does **not** mean the function f evaluated at n, but instead it is a shorthand for the function itself (leaving out domain and codomain and only giving the rule of correspondence of the function).

3. This is particularly useful if you do not want to ignore constant factors. For example the median of n elements can be determined using $\frac{3}{2}n + o(n)$ comparisons.

Asymptotic Notation

Abuse of notation

4. People write $\mathcal{O}(f(n)) = \mathcal{O}(g(n))$, when they mean $\mathcal{O}(f(n)) \subseteq \mathcal{O}(g(n))$. Again this is not an equality.

2. In this context $f(n)$ does **not** mean the function f evaluated at n, but instead it is a shorthand for the function itself (leaving out domain and codomain and only giving the rule of correspondence of the function).

3. This is particularly useful if you do not want to ignore constant factors. For example the median of n elements can be determined using $\frac{3}{2}n + o(n)$ comparisons.

Asymptotic Notation in Equations

How do we interpret an expression like:

$$2n^2 + 3n + 1 = 2n^2 + \Theta(n)$$

Here, $\Theta(n)$ stands for an anonymous function in the set $\Theta(n)$ that makes the expression true.

Note that $\Theta(n)$ is on the right hand side, otw. this interpretation is wrong.

Asymptotic Notation in Equations

How do we interpret an expression like:

$$2n^2 + \mathcal{O}(n) = \Theta(n^2)$$

Regardless of how we choose the anonymous function $f(n) \in \mathcal{O}(n)$ there is an anonymous function $g(n) \in \Theta(n^2)$ that makes the expression true.

Asymptotic Notation in Equations

The $\Theta(i)$-symbol on the left represents one anonymous function $f : \mathbb{N} \to \mathbb{R}^+$, and then $\sum_i f(i)$ is computed.

How do we interpret an expression like:

$$\sum_{i=1}^{n} \Theta(i) = \Theta(n^2)$$

Careful!

"It is understood" that every occurence of an $\mathcal{O}$-symbol (or $\Theta, \Omega, o, \omega$) on the left represents one anonymous function.

Hence, the left side is **not** equal to

$$\Theta(1) + \Theta(2) + \cdots + \Theta(n-1) + \Theta(n)$$

$\Theta(1)+\Theta(2)+\cdots+\Theta(n-1)+\Theta(n)$ does not really have a reasonable interpretation.

Asymptotic Notation in Equations

We can view an expression containing asymptotic notation as generating a set:

$$n^2 \cdot \mathcal{O}(n) + \mathcal{O}(\log n)$$

represents

$$\left\{ f : \mathbb{N} \to \mathbb{R}^+ \mid f(n) = n^2 \cdot g(n) + h(n) \right.$$
$$\left. \text{with } g(n) \in \mathcal{O}(n) \text{ and } h(n) \in \mathcal{O}(\log n) \right\}$$

Recall that according to the previous slide e.g. the expressions $\sum_{i=1}^{n} \mathcal{O}(i)$ and $\sum_{i=1}^{n/2} \mathcal{O}(i) + \sum_{i=n/2+1}^{n} \mathcal{O}(i)$ generate different sets.

Asymptotic Notation in Equations

Then an asymptotic equation can be interpreted as containement btw. two sets:

$$n^2 \cdot \mathcal{O}(n) + \mathcal{O}(\log n) = \Theta(n^2)$$

represents

$$n^2 \cdot \mathcal{O}(n) + \mathcal{O}(\log n) \subseteq \Theta(n^2)$$

Note that the equation does not hold.

Asymptotic Notation

Lemma 3
Let f, g be functions with the property $\exists n_0 > 0\,\forall n \geq n_0 : f(n) > 0$ (the same for g). Then

- $c \cdot f(n) \in \Theta(f(n))$ *for any constant* c
- $\mathcal{O}(f(n)) + \mathcal{O}(g(n)) = \mathcal{O}(f(n) + g(n))$
- $\mathcal{O}(f(n)) \cdot \mathcal{O}(g(n)) = \mathcal{O}(f(n) \cdot g(n))$
- $\mathcal{O}(f(n)) + \mathcal{O}(g(n)) = \mathcal{O}(\max\{f(n), g(n)\})$

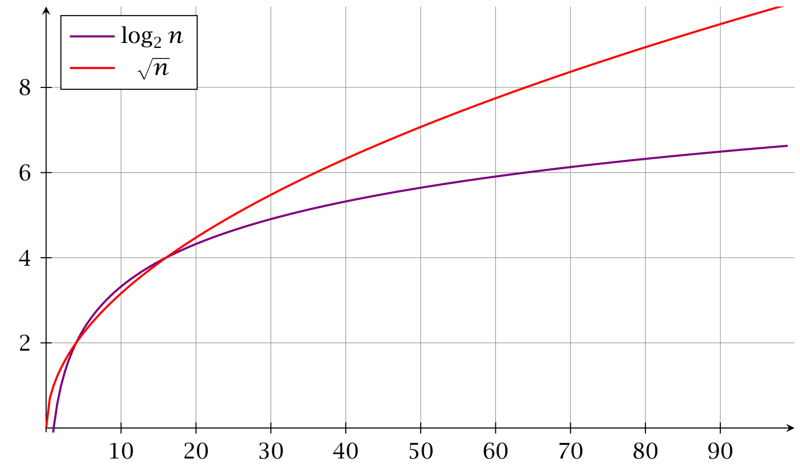
The expressions also hold for Ω. Note that this means that $f(n) + g(n) \in \Theta(\max\{f(n), g(n)\})$.

Asymptotic Notation

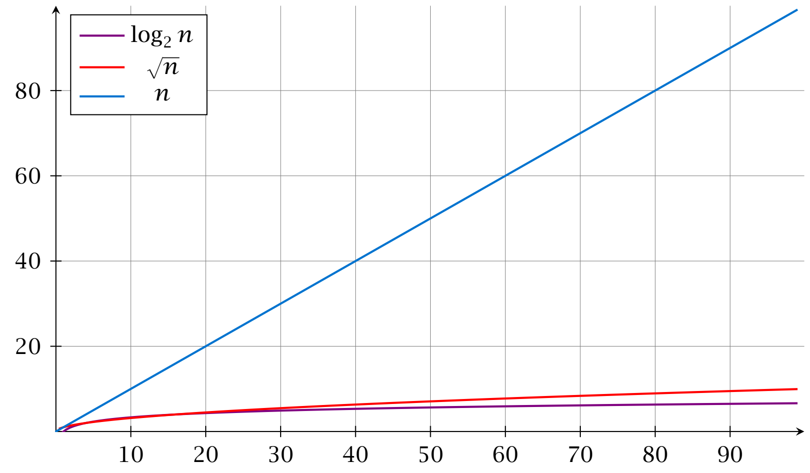
Comments

- Do not use asymptotic notation within induction proofs.
- For any constants a, b we have $\log_a n = \Theta(\log_b n)$. Therefore, we will usually ignore the base of a logarithm within asymptotic notation.
- In general $\log n = \log_2 n$, i.e., we use 2 as the default base for the logarithm.

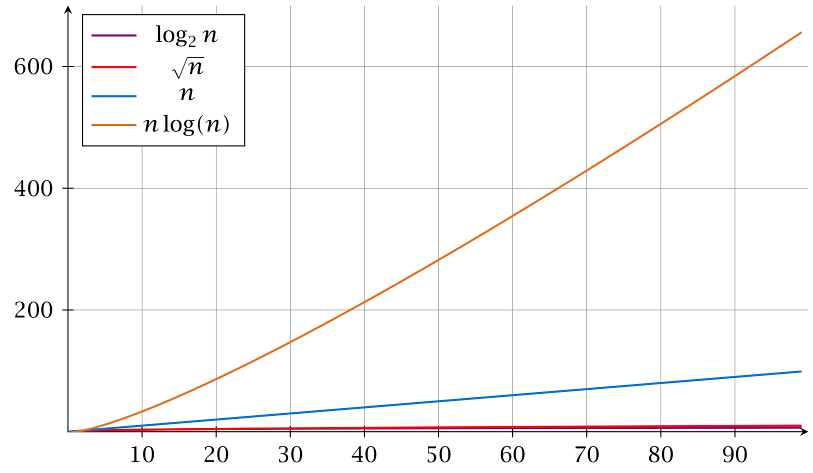
Funktionen



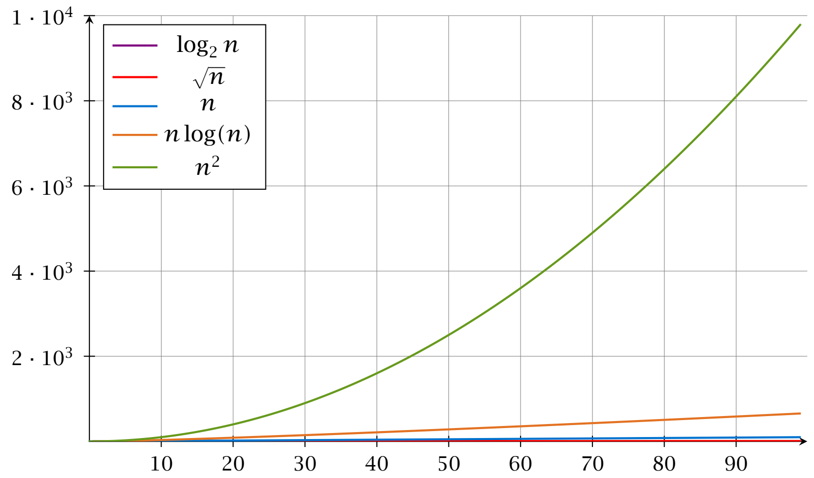
Funktionen



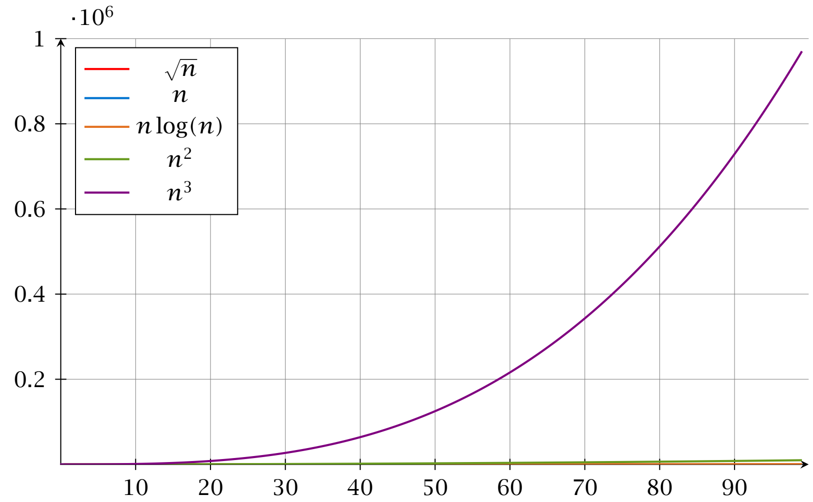
Funktionen



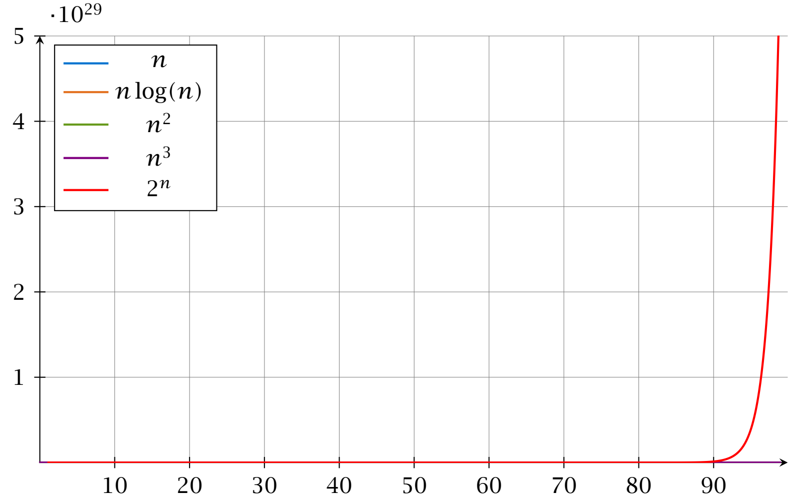
Funktionen



Funktionen



Funktionen



Laufzeiten

Funktion $f(n)$	**Eingabelänge** n: 10	10^2	10^3	10^4	10^5	10^6	10^7	10^8
$\log n$	33ns	66ns	0.1µs	0.1µs	0.2µs	0.2µs	0.2µs	0.3µs
$\sqrt{n}$	32ns	0.1µs	0.3µs	1µs	3.1µs	10µs	31µs	0.1ms
n	100ns	1µs	10µs	0.1ms	1ms	10ms	0.1s	1s
$n\log n$	0.3µs	6.6µs	0.1ms	1.3ms	16ms	0.2s	2.3s	27s
$n^{3/2}$	0.3µs	10µs	0.3ms	10ms	0.3s	10s	5.2min	2.7h
n^2	1µs	0.1ms	10ms	1s	1.7min	2.8h	11d	3.2y
n^3	10µs	10ms	10s	2.8h	115d	317y	$3.2\cdot 10^5$y	
1.1^n	26ns	0.1ms	$7.8\cdot 10^{25}$y					
2^n	10µs	$4\cdot 10^{14}$y						
$n!$	36ms	$3\cdot 10^{142}$y						

1 Operation = 10ns; 100MHz

Alter des Universums: ca. $13.8 \cdot 10^9$y

Asymptotic Notation

In general asymptotic classification of running times is a good measure for comparing algorithms:

- If the running time analysis is tight and actually occurs in practise (i.e., the asymptotic bound is not a purely theoretical worst-case bound), then the algorithm that has better asymptotic running time will always outperform a weaker algorithm for large enough values of n.
- However, suppose that I have two algorithms:
  - Algorithm A. Running time $f(n) = 1000\log n = \mathcal{O}(\log n)$.
  - Algorithm B. Running time $g(n) = \log^2 n$.

  Clearly $f = o(g)$. However, as long as $\log n \leq 1000$ Algorithm B will be more efficient.

Multiple Variables in Asymptotic Notation

Sometimes the input for an algorithm consists of several parameters (e.g., nodes and edges of a graph (n and m)).

If we want to make asympotic statements for $n \to \infty$ and $m \to \infty$ we have to extend the definition to multiple variables.

Formal Definition

Let f, g denote functions from $\mathbb{N}^d$ to $\mathbb{R}_0^+$.

- $\mathcal{O}(f) = \{g \mid \exists c > 0\ \exists N \in \mathbb{N}_0\ \forall \vec{n}$ with $n_i \geq N$ for some i :
  $[g(\vec{n}) \leq c \cdot f(\vec{n})]\}$
  (set of functions that asymptotically grow not faster than f)

Multiple Variables in Asymptotic Notation

Example 4

- $f : \mathbb{N} \to \mathbb{R}_0^+$, $f(n,m) = 1$ und $g : \mathbb{N} \to \mathbb{R}_0^+$, $g(n,m) = n - 1$
  then $f = \mathcal{O}(g)$ does not hold
- $f : \mathbb{N} \to \mathbb{R}_0^+$, $f(n,m) = 1$ und $g : \mathbb{N} \to \mathbb{R}_0^+$, $g(n,m) = n$
  then: $f = \mathcal{O}(g)$
- $f : \mathbb{N}_0 \to \mathbb{R}_0^+$, $f(n,m) = 1$ und $g : \mathbb{N}_0 \to \mathbb{R}_0^+$, $g(n,m) = n$
  then $f = \mathcal{O}(g)$ does not hold

5 Asymptotic Notation

Bibliography

[MS08] Kurt Mehlhorn, Peter Sanders:
Algorithms and Data Structures — The Basic Toolbox,
Springer, 2008

[CLRS90] Thomas H. Cormen, Charles E. Leiserson, Ron L. Rivest, Clifford Stein:
Introduction to algorithms (3rd ed.),
McGraw-Hill, 2009

Mainly Chapter 3 of [CLRS90]. [MS08] covers this topic in chapter 2.1 but not very detailed.

6 Recurrences

Algorithm 2 mergesort(list L)

```
1: n ← size(L)
2: if n ≤ 1 return L
3: L1 ← L[1 ··· ⌊n/2⌋]
4: L2 ← L[⌊n/2⌋ + 1 ··· n]
5: mergesort(L1)
6: mergesort(L2)
7: L ← merge(L1, L2)
8: return L
```

This algorithm requires

$$T(n) = T\left(\left\lceil \frac{n}{2} \right\rceil\right) + T\left(\left\lfloor \frac{n}{2} \right\rfloor\right) + \mathcal{O}(n) \le 2T\left(\left\lceil \frac{n}{2} \right\rceil\right) + \mathcal{O}(n)$$

comparisons when $n > 1$ and 0 comparisons when $n \le 1$.

Recurrences

How do we bring the expression for the number of comparisons ($\approx$ running time) into a closed form?

For this we need to solve the recurrence.

Methods for Solving Recurrences

1. **Guessing+Induction**
   Guess the right solution and prove that it is correct via induction. It needs experience to make the right guess.
2. **Master Theorem**
   For a lot of recurrences that appear in the analysis of algorithms this theorem can be used to obtain tight asymptotic bounds. It does not provide exact solutions.
3. **Characteristic Polynomial**
   Linear homogenous recurrences can be solved via this method.

Methods for Solving Recurrences

4. **Generating Functions**
   A more general technique that allows to solve certain types of linear inhomogenous relations and also sometimes non-linear recurrence relations.
5. **Transformation of the Recurrence**
   Sometimes one can transform the given recurrence relations so that it e.g. becomes linear and can therefore be solved with one of the other techniques.

6.1 Guessing+Induction

First we need to get rid of the $\mathcal{O}$-notation in our recurrence:

$$T(n) \le \begin{cases} 2T(\lceil \frac{n}{2} \rceil) + cn & n \ge 2 \\ 0 & \text{otherwise} \end{cases}$$

Informal way:
Assume that instead we have

$$T(n) \le \begin{cases} 2T(\frac{n}{2}) + cn & n \ge 2 \\ 0 & \text{otherwise} \end{cases}$$

One way of solving such a recurrence is to guess a solution, and check that it is correct by plugging it in.

6.1 Guessing+Induction

Suppose we guess $T(n) \le dn \log n$ for a constant d. Then

$$\begin{aligned} T(n) &\le 2T\Big(\frac{n}{2}\Big) + cn \\ &\le 2\Big(d\frac{n}{2}\log\frac{n}{2}\Big) + cn \\ &= dn(\log n - 1) + cn \\ &= dn\log n + (c-d)n \\ &\le dn\log n \end{aligned}$$

if we choose $d \ge c$.

Formally, this is not correct if n is not a power of 2. Also even in this case one would need to do an induction proof.

6.1 Guessing+Induction

$$T(n) \le \begin{cases} 2T(\frac{n}{2}) + cn & n \ge 16 \\ b & \text{otw.} \end{cases}$$

Guess: $T(n) \le dn\log n$.

Proof. (by induction)

- **base case** $(2 \le n < 16)$: true if we choose $d \ge b$.
- **induction step** $n/2 \to n$:

  Let $n = 2^k \ge 16$. Suppose statem. is true for $n' = n/2$. We prove it for n:

$$\begin{aligned} T(n) &\le 2T\left(\frac{n}{2}\right) + cn \\ &\le 2\left(d\frac{n}{2}\log\frac{n}{2}\right) + cn \\ &= dn(\log n - 1) + cn \\ &= dn\log n + (c-d)n \\ &\le dn\log n \end{aligned}$$

- Note that this proves the statement for $n = 2^k$, $k \in \mathbb{N}_{\ge 1}$, as the statement is wrong for $n = 1$.
- The base case is usually omitted, as it is the same for different recurrences.

Hence, statement is true if we choose $d \ge c$.

6.1 Guessing+Induction

How do we get a result for all values of n?

We consider the following recurrence instead of the original one:

$$T(n) \le \begin{cases} 2T(\lceil\frac{n}{2}\rceil) + cn & n \ge 16 \\ b & \text{otherwise} \end{cases}$$

Note that we can do this as for constant-sized inputs the running time is always some constant (b in the above case).

6.1 Guessing+Induction

We also make a guess of $T(n) \le dn\log n$ and get

$$\begin{aligned} T(n) &\le 2T\left(\left\lceil\frac{n}{2}\right\rceil\right) + cn \\ &\le 2\left(d\left\lceil\frac{n}{2}\right\rceil\log\left\lceil\frac{n}{2}\right\rceil\right) + cn \\ \boxed{\lceil\tfrac{n}{2}\rceil \le \tfrac{n}{2}+1} \quad &\le 2(d(n/2+1)\log(n/2+1)) + cn \\ \boxed{\tfrac{n}{2}+1 \le \tfrac{9}{16}n} \quad &\le dn\log\left(\frac{9}{16}n\right) + 2d\log n + cn \\ \boxed{\log\tfrac{9}{16}n = \log n + (\log 9 - 4)} \quad &= dn\log n + (\log 9 - 4)dn + 2d\log n + cn \\ \boxed{\log n \le \tfrac{n}{4}} \quad &\le dn\log n + (\log 9 - 3.5)dn + cn \\ &\le dn\log n - 0.33dn + cn \\ &\le dn\log n \end{aligned}$$

for a suitable choice of d.

6.2 Master Theorem

Note that the cases do not cover all possibilities.

Lemma 5

Let $a \ge 1, b > 1$ and $\epsilon > 0$ denote constants. Consider the recurrence

$$T(n) = aT\left(\frac{n}{b}\right) + f(n) .$$

Case 1.
If $f(n) = \mathcal{O}(n^{\log_b(a)-\epsilon})$ then $T(n) = \Theta(n^{\log_b a})$.

Case 2.
If $f(n) = \Theta(n^{\log_b(a)}\log^k n)$ then $T(n) = \Theta(n^{\log_b a}\log^{k+1} n)$, $k \ge 0$.

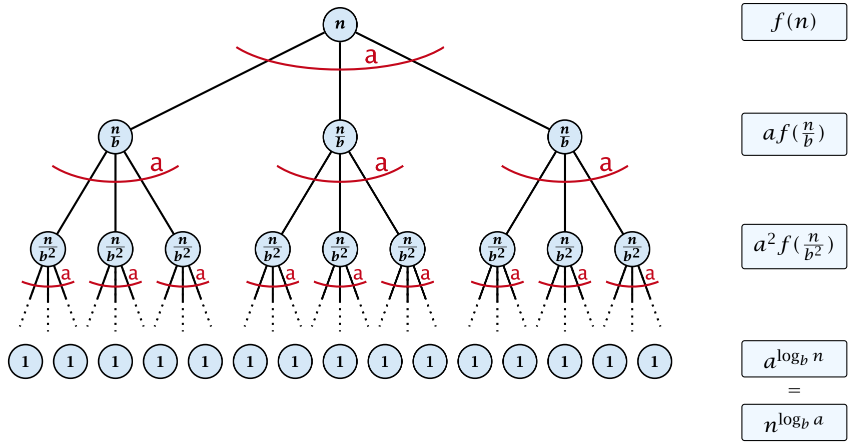
Case 3.
If $f(n) = \Omega(n^{\log_b(a)+\epsilon})$ and for sufficiently large n $af(\frac{n}{b}) \le cf(n)$ for some constant $c < 1$ then $T(n) = \Theta(f(n))$.

6.2 Master Theorem

We prove the Master Theorem for the case that n is of the form b^ℓ, and we assume that the non-recursive case occurs for problem size 1 and incurs cost 1.

The Recursion Tree

The running time of a recursive algorithm can be visualized by a recursion tree:



6.2 Master Theorem

This gives

$$T(n) = n^{\log_b a} + \sum_{i=0}^{\log_b n-1} a^i f\left(\frac{n}{b^i}\right) .$$

Case 1. Now suppose that $f(n) \le c n^{\log_b a - \epsilon}$.

$$\begin{aligned} T(n) - n^{\log_b a} &= \sum_{i=0}^{\log_b n-1} a^i f\left(\frac{n}{b^i}\right) \\ &\le c \sum_{i=0}^{\log_b n-1} a^i \left(\frac{n}{b^i}\right)^{\log_b a-\epsilon} \\ &= c n^{\log_b a-\epsilon} \sum_{i=0}^{\log_b n-1} \left(b^\epsilon\right)^i \\ &= c n^{\log_b a-\epsilon}(b^{\epsilon \log_b n} - 1)/(b^\epsilon - 1) \\ &= c n^{\log_b a-\epsilon}(n^\epsilon - 1)/(b^\epsilon - 1) \\ &= \frac{c}{b^\epsilon - 1} n^{\log_b a}(n^\epsilon - 1)/(n^\epsilon) \end{aligned}$$

$b^{-i(\log_b a - \epsilon)} = b^{\epsilon i}(b^{\log_b a})^{-i} = b^{\epsilon i} a^{-i}$

$\sum_{i=0}^k q^i = \frac{q^{k+1}-1}{q-1}$

Hence,

$$T(n) \le \left(\frac{c}{b^\epsilon - 1} + 1\right) n^{\log_b(a)}$$

$\Rightarrow T(n) = \mathcal{O}(n^{\log_b a})$.

Case 2. Now suppose that $f(n) \le cn^{\log_b a}$.

$$
\begin{aligned}
T(n) - n^{\log_b a} &= \sum_{i=0}^{\log_b n - 1} a^i f\left(\frac{n}{b^i}\right) \\
&\le c \sum_{i=0}^{\log_b n - 1} a^i \left(\frac{n}{b^i}\right)^{\log_b a} \\
&= cn^{\log_b a} \sum_{i=0}^{\log_b n - 1} 1 \\
&= cn^{\log_b a} \log_b n
\end{aligned}
$$

Hence,

$$T(n) = \mathcal{O}(n^{\log_b a} \log_b n)$$

$\Rightarrow T(n) = \mathcal{O}(n^{\log_b a} \log n)$.

Case 2. Now suppose that $f(n) \ge cn^{\log_b a}$.

$$
\begin{aligned}
T(n) - n^{\log_b a} &= \sum_{i=0}^{\log_b n - 1} a^i f\left(\frac{n}{b^i}\right) \\
&\ge c \sum_{i=0}^{\log_b n - 1} a^i \left(\frac{n}{b^i}\right)^{\log_b a} \\
&= cn^{\log_b a} \sum_{i=0}^{\log_b n - 1} 1 \\
&= cn^{\log_b a} \log_b n
\end{aligned}
$$

Hence,

$$T(n) = \Omega(n^{\log_b a} \log_b n)$$

$\Rightarrow T(n) = \Omega(n^{\log_b a} \log n)$.

Case 2. Now suppose that $f(n) \le cn^{\log_b a}(\log_b(n))^k$.

$$
\begin{aligned}
T(n) - n^{\log_b a} &= \sum_{i=0}^{\log_b n - 1} a^i f\left(\frac{n}{b^i}\right) \\
&\le c \sum_{i=0}^{\log_b n - 1} a^i \left(\frac{n}{b^i}\right)^{\log_b a} \cdot \left(\log_b\left(\frac{n}{b^i}\right)\right)^k \\
&= cn^{\log_b a} \sum_{i=0}^{\ell - 1} \left(\log_b\left(\frac{b^\ell}{b^i}\right)\right)^k \\
&= cn^{\log_b a} \sum_{i=0}^{\ell - 1} (\ell - i)^k \\
&= cn^{\log_b a} \sum_{i=1}^{\ell} i^k \quad \approx \frac{1}{k}\ell^{k+1} \\
&\approx \frac{c}{k} n^{\log_b a} \ell^{k+1}
\end{aligned}
$$

$n = b^\ell \Rightarrow \ell = \log_b n$

$\Rightarrow T(n) = \mathcal{O}(n^{\log_b a} \log^{k+1} n)$.

Case 3. Now suppose that $f(n) \ge dn^{\log_b a + \epsilon}$, and that for sufficiently large n: $af(n/b) \le cf(n)$, for $c < 1$.

From this we get $a^i f(n/b^i) \le c^i f(n)$, where we assume that $n/b^{i-1} \ge n_0$ is still sufficiently large.

$$
\begin{aligned}
T(n) - n^{\log_b a} &= \sum_{i=0}^{\log_b n - 1} a^i f\left(\frac{n}{b^i}\right) \\
&\le \sum_{i=0}^{\log_b n - 1} c^i f(n) + \mathcal{O}(n^{\log_b a}) \\
&\le \frac{1}{1-c} f(n) + \mathcal{O}(n^{\log_b a})
\end{aligned}
$$

$q < 1 : \sum_{i=0}^{n} q^i = \frac{1-q^{n+1}}{1-q} \le \frac{1}{1-q}$

Hence,

$$T(n) \le \mathcal{O}(f(n))$$

$\Rightarrow T(n) = \Theta(f(n))$.

Where did we use $f(n) \ge \Omega(n^{\log_b a + \epsilon})$?

Example: Multiplying Two Integers

Suppose we want to multiply two n-bit Integers, but our registers can only perform operations on integers of constant size.

For this we first need to be able to add two integers A and B:

```
   1 1 0 1 1 0 1 0 1    A
   1 0 0 0 1 0 0 1 1    B
  1 0 0 1 1 0 1 1 1         (carries)
 -------------------
 1 0 1 1 0 0 1 0 0 0
```

This gives that two n-bit integers can be added in time $\mathcal{O}(n)$.

Example: Multiplying Two Integers

Suppose that we want to multiply an n-bit integer A and an m-bit integer B ($m \le n$).

```
        1 0 0 0 1 × 1 0 1 1
       ---------------------
                  1 0 0 0 1
                1 0 0 0 1 0
              0 0 0 0 0 0 0
            1 0 0 0 1 0 0 0
       ---------------------
            1 0 1 1 1 0 1 1
```

- This is also nown as the "school method" for multiplying integers.
- Note that the intermediate numbers that are generated can have at most $m + n \le 2n$ bits.

Time requirement:

- Computing intermediate results: $\mathcal{O}(nm)$.
- Adding m numbers of length $\le 2n$: $\mathcal{O}((m+n)m) = \mathcal{O}(nm)$.

Example: Multiplying Two Integers

A recursive approach:

Suppose that integers A and B are of length $n = 2^k$, for some k.

B_1	B_0	×	A_1	A_0

Then it holds that

$$A = A_1 \cdot 2^{\frac{n}{2}} + A_0 \text{ and } B = B_1 \cdot 2^{\frac{n}{2}} + B_0$$

Hence,

$$A \cdot B = A_1 B_1 \cdot 2^n + (A_1 B_0 + A_0 B_1) \cdot 2^{\frac{n}{2}} + A_0 B_0$$

Example: Multiplying Two Integers

Algorithm 3 mult(A, B)	
if $\lvert A\rvert = \lvert B\rvert = 1$ **then**	$\mathcal{O}(1)$
return $a_0 \cdot b_0$	$\mathcal{O}(1)$
split A into A_0 and A_1	$\mathcal{O}(n)$
split B into B_0 and B_1	$\mathcal{O}(n)$
$Z_2 \leftarrow \text{mult}(A_1, B_1)$	$T(\frac{n}{2})$
$Z_1 \leftarrow \text{mult}(A_1, B_0) + \text{mult}(A_0, B_1)$	$2T(\frac{n}{2}) + \mathcal{O}(n)$
$Z_0 \leftarrow \text{mult}(A_0, B_0)$	$T(\frac{n}{2})$
return $Z_2 \cdot 2^n + Z_1 \cdot 2^{\frac{n}{2}} + Z_0$	$\mathcal{O}(n)$

We get the following recurrence:

$$T(n) = 4T\left(\frac{n}{2}\right) + \mathcal{O}(n) \ .$$

Example: Multiplying Two Integers

Master Theorem: Recurrence: $T[n] = aT(\frac{n}{b}) + f(n)$.

- Case 1: $f(n) = \mathcal{O}(n^{\log_b a - \epsilon})$ $\quad T(n) = \Theta(n^{\log_b a})$
- Case 2: $f(n) = \Theta(n^{\log_b a}\log^k n)$ $\quad T(n) = \Theta(n^{\log_b a}\log^{k+1} n)$
- Case 3: $f(n) = \Omega(n^{\log_b a + \epsilon})$ $\quad T(n) = \Theta(f(n))$

In our case $a = 4$, $b = 2$, and $f(n) = \Theta(n)$. Hence, we are in Case 1, since $n = \mathcal{O}(n^{2-\epsilon}) = \mathcal{O}(n^{\log_b a - \epsilon})$.

We get a running time of $\mathcal{O}(n^2)$ for our algorithm.

$\Longrightarrow$ Not better then the "school method".

Example: Multiplying Two Integers

We can use the following identity to compute Z_1:

$$\begin{aligned} Z_1 &= A_1B_0 + A_0B_1 \\ &= (A_0 + A_1)\cdot(B_0 + B_1) - \overbrace{A_1B_1}^{=Z_2} - \overbrace{A_0B_0}^{=Z_0} \end{aligned}$$

Hence,

Algorithm 4 mult(A, B)

1:	**if** $\lvert A\rvert = \lvert B\rvert = 1$ **then**	$\mathcal{O}(1)$
2:	**return** $a_0 \cdot b_0$	$\mathcal{O}(1)$
3:	split A into A_0 and A_1	$\mathcal{O}(n)$
4:	split B into B_0 and B_1	$\mathcal{O}(n)$
5:	$Z_2 \leftarrow \mathrm{mult}(A_1, B_1)$	$T(\frac{n}{2})$
6:	$Z_0 \leftarrow \mathrm{mult}(A_0, B_0)$	$T(\frac{n}{2})$
7:	$Z_1 \leftarrow \mathrm{mult}(A_0 + A_1, B_0 + B_1) - Z_2 - Z_0$	$T(\frac{n}{2}) + \mathcal{O}(n)$
8:	**return** $Z_2 \cdot 2^n + Z_1 \cdot 2^{\frac{n}{2}} + Z_0$	$\mathcal{O}(n)$

A more precise (correct) analysis would say that computing Z_1 needs time $T(\frac{n}{2} + 1) + \mathcal{O}(n)$.

Example: Multiplying Two Integers

We get the following recurrence:

$$T(n) = 3T\left(\frac{n}{2}\right) + \mathcal{O}(n) \ .$$

Master Theorem: Recurrence: $T[n] = aT(\frac{n}{b}) + f(n)$.

- Case 1: $f(n) = \mathcal{O}(n^{\log_b a - \epsilon})$ $\quad T(n) = \Theta(n^{\log_b a})$
- Case 2: $f(n) = \Theta(n^{\log_b a}\log^k n)$ $\quad T(n) = \Theta(n^{\log_b a}\log^{k+1} n)$
- Case 3: $f(n) = \Omega(n^{\log_b a + \epsilon})$ $\quad T(n) = \Theta(f(n))$

Again we are in Case 1. We get a running time of $\Theta(n^{\log_2 3}) \approx \Theta(n^{1.59})$.

A huge improvement over the "school method".

6.3 The Characteristic Polynomial

Consider the recurrence relation:

$$c_0T(n) + c_1T(n-1) + c_2T(n-2) + \cdots + c_kT(n-k) = f(n)$$

This is the general form of a linear recurrence relation of order k with constant coefficients ($c_0, c_k \neq 0$).

- $T(n)$ only depends on the k preceding values. This means the recurrence relation is of order k.
- The recurrence is linear as there are no products of $T[n]$'s.
- If $f(n) = 0$ then the recurrence relation becomes a linear, homogenous recurrence relation of order k.

Note that we ignore boundary conditions for the moment.

6.3 The Characteristic Polynomial

Observations:

- The solution $T[1], T[2], T[3], \ldots$ is completely determined by a set of boundary conditions that specify values for $T[1], \ldots, T[k]$.
- In fact, any k consecutive values completely determine the solution.
- k non-concecutive values might not be an appropriate set of boundary conditions (depends on the problem).

Approach:

- First determine all solutions that satisfy recurrence relation.
- Then pick the right one by analyzing boundary conditions.
- First consider the homogenous case.

The Homogenous Case

The solution space

$$S = \left\{ \mathcal{T} = T[1], T[2], T[3], \ldots \mid \mathcal{T} \text{ fulfills recurrence relation} \right\}$$

is a vector space. This means that if $\mathcal{T}_1, \mathcal{T}_2 \in S$, then also $\alpha\mathcal{T}_1 + \beta\mathcal{T}_2 \in S$, for arbitrary constants α, β.

How do we find a non-trivial solution?

We guess that the solution is of the form λ^n, $\lambda \neq 0$, and see what happens. In order for this guess to fulfill the recurrence we need

$$c_0\lambda^n + c_1\lambda^{n-1} + c_2 \cdot \lambda^{n-2} + \cdots + c_k \cdot \lambda^{n-k} = 0$$

for all $n \geq k$.

The Homogenous Case

Dividing by λ^{n-k} gives that all these constraints are identical to

$$\underbrace{c_0\lambda^k + c_1\lambda^{k-1} + c_2 \cdot \lambda^{k-2} + \cdots + c_k}_{\text{characteristic polynomial } P[\lambda]} = 0$$

This means that if λ_i is a root (Nullstelle) of $P[\lambda]$ then $T[n] = \lambda_i^n$ is a solution to the recurrence relation.

Let $\lambda_1, \ldots, \lambda_k$ be the k (complex) roots of $P[\lambda]$. Then, because of the vector space property

$$\alpha_1\lambda_1^n + \alpha_2\lambda_2^n + \cdots + \alpha_k\lambda_k^n$$

is a solution for arbitrary values α_i.

The Homogenous Case

Lemma 6

Assume that the characteristic polynomial has k distinct roots $\lambda_1, \ldots, \lambda_k$. Then all solutions to the recurrence relation are of the form

$$\alpha_1\lambda_1^n + \alpha_2\lambda_2^n + \cdots + \alpha_k\lambda_k^n \ .$$

Proof.

There is one solution for every possible choice of boundary conditions for $T[1], \ldots, T[k]$.

We show that the above set of solutions contains one solution for every choice of boundary conditions.

The Homogenous Case

Proof (cont.).

Suppose I am given boundary conditions $T[i]$ and I want to see whether I can choose the $\alpha_i's$ such that these conditions are met:

$$\begin{array}{ccccccccc}\alpha_1\cdot\lambda_1 & + & \alpha_2\cdot\lambda_2 & + & \cdots & + & \alpha_k\cdot\lambda_k & = & T[1]\\ \alpha_1\cdot\lambda_1^2 & + & \alpha_2\cdot\lambda_2^2 & + & \cdots & + & \alpha_k\cdot\lambda_k^2 & = & T[2]\\ & & & & \vdots & & & & \\ \alpha_1\cdot\lambda_1^k & + & \alpha_2\cdot\lambda_2^k & + & \cdots & + & \alpha_k\cdot\lambda_k^k & = & T[k]\end{array}$$

The Homogenous Case

Proof (cont.).

Suppose I am given boundary conditions $T[i]$ and I want to see whether I can choose the $\alpha_i's$ such that these conditions are met:

$$\begin{pmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_k\\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_k^2\\ & & \vdots & \\ \lambda_1^k & \lambda_2^k & \cdots & \lambda_k^k\end{pmatrix}\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_k\end{pmatrix}=\begin{pmatrix}T[1]\\ T[2]\\ \vdots\\ T[k]\end{pmatrix}$$

We show that the column vectors are linearly independent. Then the above equation has a solution.

Computing the Determinant

$$\begin{vmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_{k-1} & \lambda_k\\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_{k-1}^2 & \lambda_k^2\\ \vdots & \vdots & & \vdots & \vdots\\ \lambda_1^k & \lambda_2^k & \cdots & \lambda_{k-1}^k & \lambda_k^k\end{vmatrix}=\prod_{i=1}^k\lambda_i\cdot\begin{vmatrix}1 & 1 & \cdots & 1 & 1\\ \lambda_1 & \lambda_2 & \cdots & \lambda_{k-1} & \lambda_k\\ \vdots & \vdots & & \vdots & \vdots\\ \lambda_1^{k-1} & \lambda_2^{k-1} & \cdots & \lambda_{k-1}^{k-1} & \lambda_k^{k-1}\end{vmatrix}$$

$$=\prod_{i=1}^k\lambda_i\cdot\begin{vmatrix}1 & \lambda_1 & \cdots & \lambda_1^{k-2} & \lambda_1^{k-1}\\ 1 & \lambda_2 & \cdots & \lambda_2^{k-2} & \lambda_2^{k-1}\\ \vdots & \vdots & & \vdots & \vdots\\ 1 & \lambda_k & \cdots & \lambda_k^{k-2} & \lambda_k^{k-1}\end{vmatrix}$$

Computing the Determinant

$$\begin{vmatrix}1 & \lambda_1 & \cdots & \lambda_1^{k-2} & \lambda_1^{k-1}\\ 1 & \lambda_2 & \cdots & \lambda_2^{k-2} & \lambda_2^{k-1}\\ \vdots & \vdots & & \vdots & \vdots\\ 1 & \lambda_k & \cdots & \lambda_k^{k-2} & \lambda_k^{k-1}\end{vmatrix}=$$

$$\begin{vmatrix}1 & \lambda_1-\lambda_1\cdot 1 & \cdots & \lambda_1^{k-2}-\lambda_1\cdot\lambda_1^{k-3} & \lambda_1^{k-1}-\lambda_1\cdot\lambda_1^{k-2}\\ 1 & \lambda_2-\lambda_1\cdot 1 & \cdots & \lambda_2^{k-2}-\lambda_1\cdot\lambda_2^{k-3} & \lambda_2^{k-1}-\lambda_1\cdot\lambda_2^{k-2}\\ \vdots & \vdots & & \vdots & \vdots\\ 1 & \lambda_k-\lambda_1\cdot 1 & \cdots & \lambda_k^{k-2}-\lambda_1\cdot\lambda_k^{k-3} & \lambda_k^{k-1}-\lambda_1\cdot\lambda_k^{k-2}\end{vmatrix}$$

Computing the Determinant

$$\begin{vmatrix} 1 & \lambda_1-\lambda_1\cdot 1 & \cdots & \lambda_1^{k-2}-\lambda_1\cdot\lambda_1^{k-3} & \lambda_1^{k-1}-\lambda_1\cdot\lambda_1^{k-2} \\ 1 & \lambda_2-\lambda_1\cdot 1 & \cdots & \lambda_2^{k-2}-\lambda_1\cdot\lambda_2^{k-3} & \lambda_2^{k-1}-\lambda_1\cdot\lambda_2^{k-2} \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & \lambda_k-\lambda_1\cdot 1 & \cdots & \lambda_k^{k-2}-\lambda_1\cdot\lambda_k^{k-3} & \lambda_k^{k-1}-\lambda_1\cdot\lambda_k^{k-2} \end{vmatrix} =$$

$$\begin{vmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & (\lambda_2-\lambda_1)\cdot 1 & \cdots & (\lambda_2-\lambda_1)\cdot\lambda_2^{k-3} & (\lambda_2-\lambda_1)\cdot\lambda_2^{k-2} \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & (\lambda_k-\lambda_1)\cdot 1 & \cdots & (\lambda_k-\lambda_1)\cdot\lambda_k^{k-3} & (\lambda_k-\lambda_1)\cdot\lambda_k^{k-2} \end{vmatrix}$$

Computing the Determinant

$$\begin{vmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & (\lambda_2-\lambda_1)\cdot 1 & \cdots & (\lambda_2-\lambda_1)\cdot\lambda_2^{k-3} & (\lambda_2-\lambda_1)\cdot\lambda_2^{k-2} \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & (\lambda_k-\lambda_1)\cdot 1 & \cdots & (\lambda_k-\lambda_1)\cdot\lambda_k^{k-3} & (\lambda_k-\lambda_1)\cdot\lambda_k^{k-2} \end{vmatrix} =$$

$$\prod_{i=2}^{k}(\lambda_i-\lambda_1)\cdot\begin{vmatrix} 1 & \lambda_2 & \cdots & \lambda_2^{k-3} & \lambda_2^{k-2} \\ \vdots & \vdots & & \vdots & \vdots \\ 1 & \lambda_k & \cdots & \lambda_k^{k-3} & \lambda_k^{k-2} \end{vmatrix}$$

Computing the Determinant

Repeating the above steps gives:

$$\begin{vmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_{k-1} & \lambda_k \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_{k-1}^2 & \lambda_k^2 \\ \vdots & \vdots & & \vdots & \vdots \\ \lambda_1^k & \lambda_2^k & \cdots & \lambda_{k-1}^k & \lambda_k^k \end{vmatrix} = \prod_{i=1}^{k}\lambda_i\cdot\prod_{i>\ell}(\lambda_i-\lambda_\ell)$$

Hence, if all λ_i's are different, then the determinant is non-zero.

The Homogeneous Case

What happens if the roots are not all distinct?

Suppose we have a root λ_i with multiplicity (Vielfachheit) at least 2. Then not only is λ_i^n a solution to the recurrence but also $n\lambda_i^n$.

To see this consider the polynomial

$$P[\lambda]\cdot\lambda^{n-k} = c_0\lambda^n + c_1\lambda^{n-1} + c_2\lambda^{n-2} + \cdots + c_k\lambda^{n-k}$$

Since λ_i is a root we can write this as $Q[\lambda]\cdot(\lambda-\lambda_i)^2$. Calculating the derivative gives a polynomial that still has root λ_i.

This means

$$c_0 n\lambda_i^{n-1} + c_1(n-1)\lambda_i^{n-2} + \cdots + c_k(n-k)\lambda_i^{n-k-1} = 0$$

Hence,

$$\underbrace{c_0 n\lambda_i^{n}}_{T[n]} + c_1\underbrace{(n-1)\lambda_i^{n-1}}_{T[n-1]} + \cdots + c_k\underbrace{(n-k)\lambda_i^{n-k}}_{T[n-k]} = 0$$

The Homogeneous Case

Suppose λ_i has multiplicity j. We know that

$$c_0 n\lambda_i^{n} + c_1(n-1)\lambda_i^{n-1} + \cdots + c_k(n-k)\lambda_i^{n-k} = 0$$

(after taking the derivative; multiplying with λ; plugging in λ_i)

Doing this again gives

$$c_0 n^2\lambda_i^{n} + c_1(n-1)^2\lambda_i^{n-1} + \cdots + c_k(n-k)^2\lambda_i^{n-k} = 0$$

We can continue $j-1$ times.

Hence, $n^\ell\lambda_i^n$ is a solution for $\ell \in 0,\dots,j-1$.

The Homogeneous Case

Lemma 7

Let $P[\lambda]$ denote the characteristic polynomial to the recurrence

$$c_0 T[n] + c_1 T[n-1] + \cdots + c_k T[n-k] = 0$$

Let λ_i, $i = 1,\dots,m$ be the (complex) roots of $P[\lambda]$ with multiplicities ℓ_i. Then the general solution to the recurrence is given by

$$T[n] = \sum_{i=1}^{m}\sum_{j=0}^{\ell_i-1} \alpha_{ij}\cdot(n^j\lambda_i^n)\ .$$

The full proof is omitted. We have only shown that any choice of α_{ij}'s is a solution to the recurrence.

Example: Fibonacci Sequence

$$T[0] = 0$$
$$T[1] = 1$$
$$T[n] = T[n-1] + T[n-2] \text{ for } n \geq 2$$

The characteristic polynomial is

$$\lambda^2 - \lambda - 1$$

Finding the roots, gives

$$\lambda_{1/2} = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 1} = \frac{1}{2}\left(1 \pm \sqrt{5}\right)$$

Example: Fibonacci Sequence

Hence, the solution is of the form

$$\alpha\left(\frac{1+\sqrt{5}}{2}\right)^n+\beta\left(\frac{1-\sqrt{5}}{2}\right)^n$$

$T[0]=0$ gives $\alpha+\beta=0$.

$T[1]=1$ gives

$$\alpha\left(\frac{1+\sqrt{5}}{2}\right)+\beta\left(\frac{1-\sqrt{5}}{2}\right)=1 \Longrightarrow \alpha-\beta=\frac{2}{\sqrt{5}}$$

Example: Fibonacci Sequence

Hence, the solution is

$$\frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n-\left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

The Inhomogeneous Case

Consider the recurrence relation:

$$c_0T(n)+c_1T(n-1)+c_2T(n-2)+\cdots+c_kT(n-k)=f(n)$$

with $f(n)\neq 0$.

While we have a fairly general technique for solving homogeneous, linear recurrence relations the inhomogeneous case is different.

The Inhomogeneous Case

The general solution of the recurrence relation is

$$T(n)=T_h(n)+T_p(n)\ ,$$

where T_h is any solution to the homogeneous equation, and T_p is one particular solution to the inhomogeneous equation.

There is no general method to find a particular solution.

The Inhomogeneous Case

Example:

$$T[n] = T[n-1] + 1 \qquad T[0] = 1$$

Then,

$$T[n-1] = T[n-2] + 1 \qquad (n \geq 2)$$

Subtracting the first from the second equation gives,

$$T[n] - T[n-1] = T[n-1] - T[n-2] \qquad (n \geq 2)$$

or

$$T[n] = 2T[n-1] - T[n-2] \qquad (n \geq 2)$$

I get a completely determined recurrence if I add $T[0] = 1$ and $T[1] = 2$.

The Inhomogeneous Case

Example: Characteristic polynomial:

$$\underbrace{\lambda^2 - 2\lambda + 1}_{(\lambda-1)^2} = 0$$

Then the solution is of the form

$$T[n] = \alpha 1^n + \beta n 1^n = \alpha + \beta n$$

$T[0] = 1$ gives $\alpha = 1$.

$T[1] = 2$ gives $1 + \beta = 2 \Longrightarrow \beta = 1$.

The Inhomogeneous Case

If $f(n)$ is a polynomial of degree r this method can be applied $r+1$ times to obtain a homogeneous equation:

$$T[n] = T[n-1] + n^2$$

Shift:

$$T[n-1] = T[n-2] + (n-1)^2 = T[n-2] + n^2 - 2n + 1$$

Difference:

$$T[n] - T[n-1] = T[n-1] - T[n-2] + 2n - 1$$

$$T[n] = 2T[n-1] - T[n-2] + 2n - 1$$

$$T[n] = 2T[n-1] - T[n-2] + 2n - 1$$

Shift:

$$\begin{aligned} T[n-1] &= 2T[n-2] - T[n-3] + 2(n-1) - 1 \\ &= 2T[n-2] - T[n-3] + 2n - 3 \end{aligned}$$

Difference:

$$\begin{aligned} T[n] - T[n-1] =& 2T[n-1] - T[n-2] + 2n - 1 \\ & - 2T[n-2] + T[n-3] - 2n + 3 \end{aligned}$$

$$T[n] = 3T[n-1] - 3T[n-2] + T[n-3] + 2$$

and so on...

6.4 Generating Functions

Definition 8 (Generating Function)

Let $(a_n)_{n\ge 0}$ be a sequence. The corresponding

- generating function (Erzeugendenfunktion) is
$$F(z) := \sum_{n\ge 0} a_n z^n \ ;$$
- exponential generating function (exponentielle Erzeugendenfunktion) is
$$F(z) := \sum_{n\ge 0} \frac{a_n}{n!} z^n \ .$$

6.4 Generating Functions

Example 9

1. The generating function of the sequence $(1, 0, 0, \ldots)$ is
$$F(z) = 1 \, .$$
2. The generating function of the sequence $(1, 1, 1, \ldots)$ is
$$F(z) = \frac{1}{1-z} \, .$$

6.4 Generating Functions

There are two different views:

A generating function is a formal power series (formale Potenzreihe).

Then the generating function is an algebraic object.

Let $f = \sum_{n\ge 0} a_n z^n$ and $g = \sum_{n\ge 0} b_n z^n$.

- **Equality:** f and g are equal if $a_n = b_n$ for all n.
- **Addition:** $f + g := \sum_{n\ge 0}(a_n + b_n)z^n$.
- **Multiplication:** $f \cdot g := \sum_{n\ge 0} c_n z^n$ with $c_n = \sum_{p=0}^{n} a_p b_{n-p}$.

There are no convergence issues here.

6.4 Generating Functions

The arithmetic view:

We view a power series as a function $f : \mathbb{C} \to \mathbb{C}$.

Then, it is important to think about convergence/convergence radius etc.

6.4 Generating Functions

What does $\sum_{n\geq 0} z^n = \frac{1}{1-z}$ mean in the algebraic view?

It means that the power series $1-z$ and the power series $\sum_{n\geq 0} z^n$ are invers, i.e.,

$$\left(1-z\right)\cdot\Big(\sum_{n\geq 0}^{\infty} z^n\Big) = 1\enspace.$$

This is well-defined.

6.4 Generating Functions

Formally the derivative of a formal power series $\sum_{n\geq 0} a_n z^n$ is defined as $\sum_{n\geq 0} n a_n z^{n-1}$.

The known rules for differentiation work for this definition. In particular, e.g. the derivative of $\frac{1}{1-z}$ is $\frac{1}{(1-z)^2}$.

Note that this requires a proof if we consider power series as algebraic objects. However, we did not prove this in the lecture.

Suppose we are given the generating function

$$\sum_{n\geq 0} z^n = \frac{1}{1-z}\enspace.$$

We can compute the derivative:

$$\underbrace{\sum_{n\geq 1} n z^{n-1}}_{\sum_{n\geq 0}(n+1)z^n} = \frac{1}{(1-z)^2}$$

Hence, the generating function of the sequence $a_n = n+1$ is $1/(1-z)^2$.

6.4 Generating Functions

We can repeat this

$$\sum_{n\geq 0}(n+1)z^n = \frac{1}{(1-z)^2}\enspace.$$

Derivative:

$$\underbrace{\sum_{n\geq 1} n(n+1)z^{n-1}}_{\sum_{n\geq 0}(n+1)(n+2)z^n} = \frac{2}{(1-z)^3}$$

Hence, the generating function of the sequence $a_n = (n+1)(n+2)$ is $\frac{2}{(1-z)^3}$.

6.4 Generating Functions

Computing the k-th derivative of $\sum z^n$.

$$\begin{aligned}\sum_{n\geq k} n(n-1)\cdot\ldots\cdot(n-k+1)z^{n-k} &= \sum_{n\geq 0}(n+k)\cdot\ldots\cdot(n+1)z^n\\ &= \frac{k!}{(1-z)^{k+1}}\enspace.\end{aligned}$$

Hence:

$$\sum_{n\geq 0}\binom{n+k}{k} z^n = \frac{1}{(1-z)^{k+1}}\enspace.$$

The generating function of the sequence $a_n = \binom{n+k}{k}$ is $\frac{1}{(1-z)^{k+1}}$.

6.4 Generating Functions

$$\sum_{n\ge 0} nz^n = \sum_{n\ge 0}(n+1)z^n - \sum_{n\ge 0} z^n = \frac{1}{(1-z)^2} - \frac{1}{1-z} = \frac{z}{(1-z)^2}$$

The generating function of the sequence $a_n = n$ is $\frac{z}{(1-z)^2}$.

6.4 Generating Functions

We know

$$\sum_{n\ge 0} y^n = \frac{1}{1-y}$$

Hence,

$$\sum_{n\ge 0} a^n z^n = \frac{1}{1-az}$$

The generating function of the sequence $f_n = a^n$ is $\frac{1}{1-az}$.

Example: $a_n = a_{n-1} + 1$, $a_0 = 1$

Suppose we have the recurrence $a_n = a_{n-1} + 1$ for $n \ge 1$ and $a_0 = 1$.

$$\begin{aligned} A(z) &= \sum_{n\ge 0} a_n z^n \\ &= a_0 + \sum_{n\ge 1}(a_{n-1}+1)z^n \\ &= 1 + z\sum_{n\ge 1} a_{n-1}z^{n-1} + \sum_{n\ge 1} z^n \\ &= z\sum_{n\ge 0} a_n z^n + \sum_{n\ge 0} z^n \\ &= zA(z) + \sum_{n\ge 0} z^n \\ &= zA(z) + \frac{1}{1-z} \end{aligned}$$

Example: $a_n = a_{n-1} + 1$, $a_0 = 1$

Solving for $A(z)$ gives

$$\sum_{n\ge 0} a_n z^n = A(z) = \frac{1}{(1-z)^2} = \sum_{n\ge 0}(n+1)z^n$$

Hence, $a_n = n + 1$.

Some Generating Functions

n-th sequence element	*generating function*
1	$\frac{1}{1-z}$
$n+1$	$\frac{1}{(1-z)^2}$
$\binom{n+k}{k}$	$\frac{1}{(1-z)^{k+1}}$
n	$\frac{z}{(1-z)^2}$
a^n	$\frac{1}{1-az}$
n^2	$\frac{z(1+z)}{(1-z)^3}$
$\frac{1}{n!}$	e^z

Some Generating Functions

n-th sequence element	*generating function*
cf_n	cF
f_n+g_n	$F+G$
$\sum_{i=0}^n f_ig_{n-i}$	$F\cdot G$
$f_{n-k}\ (n\ge k);\ \ 0$ otw.	z^kF
$\sum_{i=0}^n f_i$	$\frac{F(z)}{1-z}$
nf_n	$z\frac{\mathrm{d}F(z)}{\mathrm{d}z}$
c^nf_n	$F(cz)$

Solving Recursions with Generating Functions

1. Set $A(z)=\sum_{n\ge0}a_nz^n$.
2. Transform the right hand side so that boundary condition and recurrence relation can be plugged in.
3. Do further transformations so that the infinite sums on the right hand side can be replaced by $A(z)$.
4. Solving for $A(z)$ gives an equation of the form $A(z)=f(z)$, where hopefully $f(z)$ is a simple function.
5. Write $f(z)$ as a formal power series. Techniques:
   - partial fraction decomposition (Partialbruchzerlegung)
   - lookup in tables
6. The coefficients of the resulting power series are the a_n.

Example: $a_n=2a_{n-1}$, $a_0=1$

1. Set up generating function:

$$A(z)=\sum_{n\ge0}a_nz^n$$

2. Transform right hand side so that recurrence can be plugged in:

$$A(z)=a_0+\sum_{n\ge1}a_nz^n$$

2. Plug in:

$$A(z)=1+\sum_{n\ge1}(2a_{n-1})z^n$$

Example: $a_n = 2a_{n-1}$, $a_0 = 1$

3. Transform right hand side so that infinite sums can be replaced by $A(z)$ or by simple function.

$$\begin{aligned} A(z) &= 1 + \sum_{n\geq 1}(2a_{n-1})z^n \\ &= 1 + 2z\sum_{n\geq 1}a_{n-1}z^{n-1} \\ &= 1 + 2z\sum_{n\geq 0}a_nz^n \\ &= 1 + 2z\cdot A(z) \end{aligned}$$

4. Solve for $A(z)$.

$$A(z) = \frac{1}{1-2z}$$

Example: $a_n = 2a_{n-1}$, $a_0 = 1$

5. Rewrite $f(z)$ as a power series:

$$\sum_{n\geq 0}a_nz^n = A(z) = \frac{1}{1-2z} = \sum_{n\geq 0}2^nz^n$$

Example: $a_n = 3a_{n-1} + n$, $a_0 = 1$

1. Set up generating function:

$$A(z) = \sum_{n\geq 0}a_nz^n$$

Example: $a_n = 3a_{n-1} + n$, $a_0 = 1$

2./3. Transform right hand side:

$$\begin{aligned} A(z) &= \sum_{n\geq 0}a_nz^n \\ &= a_0 + \sum_{n\geq 1}a_nz^n \\ &= 1 + \sum_{n\geq 1}(3a_{n-1}+n)z^n \\ &= 1 + 3z\sum_{n\geq 1}a_{n-1}z^{n-1} + \sum_{n\geq 1}nz^n \\ &= 1 + 3z\sum_{n\geq 0}a_nz^n + \sum_{n\geq 0}nz^n \\ &= 1 + 3zA(z) + \frac{z}{(1-z)^2} \end{aligned}$$

Example: $a_n = 3a_{n-1} + n, a_0 = 1$

4. Solve for $A(z)$:

$$A(z) = 1 + 3zA(z) + \frac{z}{(1-z)^2}$$

gives

$$A(z) = \frac{(1-z)^2 + z}{(1-3z)(1-z)^2} = \frac{z^2 - z + 1}{(1-3z)(1-z)^2}$$

Example: $a_n = 3a_{n-1} + n, a_0 = 1$

5. Write $f(z)$ as a formal power series:

We use partial fraction decomposition:

$$\frac{z^2 - z + 1}{(1-3z)(1-z)^2} \stackrel{!}{=} \frac{A}{1-3z} + \frac{B}{1-z} + \frac{C}{(1-z)^2}$$

This gives

$$\begin{aligned} z^2 - z + 1 &= A(1-z)^2 + B(1-3z)(1-z) + C(1-3z) \\ &= A(1-2z+z^2) + B(1-4z+3z^2) + C(1-3z) \\ &= (A+3B)z^2 + (-2A-4B-3C)z + (A+B+C) \end{aligned}$$

Example: $a_n = 3a_{n-1} + n, a_0 = 1$

5. Write $f(z)$ as a formal power series:

This leads to the following conditions:

$$\begin{aligned} A + B + C &= 1 \\ 2A + 4B + 3C &= 1 \\ A + 3B &= 1 \end{aligned}$$

which gives

$$A = \frac{7}{4} \quad B = -\frac{1}{4} \quad C = -\frac{1}{2}$$

Example: $a_n = 3a_{n-1} + n, a_0 = 1$

5. Write $f(z)$ as a formal power series:

$$\begin{aligned} A(z) &= \frac{7}{4} \cdot \frac{1}{1-3z} - \frac{1}{4} \cdot \frac{1}{1-z} - \frac{1}{2} \cdot \frac{1}{(1-z)^2} \\ &= \frac{7}{4} \cdot \sum_{n \geq 0} 3^n z^n - \frac{1}{4} \cdot \sum_{n \geq 0} z^n - \frac{1}{2} \cdot \sum_{n \geq 0} (n+1) z^n \\ &= \sum_{n \geq 0} \left(\frac{7}{4} \cdot 3^n - \frac{1}{4} - \frac{1}{2}(n+1) \right) z^n \\ &= \sum_{n \geq 0} \left(\frac{7}{4} \cdot 3^n - \frac{1}{2} n - \frac{3}{4} \right) z^n \end{aligned}$$

6. This means $a_n = \frac{7}{4} 3^n - \frac{1}{2} n - \frac{3}{4}$.

6.5 Transformation of the Recurrence

Example 10

$$f_0 = 1$$
$$f_1 = 2$$
$$f_n = f_{n-1} \cdot f_{n-2} \text{ for } n \geq 2 \; .$$

Define

$$g_n := \log f_n \; .$$

Then

$$g_n = g_{n-1} + g_{n-2} \text{ for } n \geq 2$$
$$g_1 = \log 2 = 1 (\text{for } \log = \log_2),\;\; g_0 = 0$$
$$g_n = F_n \;(n\text{-th Fibonacci number})$$
$$f_n = 2^{F_n}$$

6.5 Transformation of the Recurrence

Example 11

$$f_1 = 1$$
$$f_n = 3f_{\frac{n}{2}} + n; \text{ for } n = 2^k,\ k \geq 1 \;;$$

Define

$$g_k := f_{2^k} \; .$$

Then:

$$g_0 = 1$$
$$g_k = 3g_{k-1} + 2^k,\;\; k \geq 1$$

6 Recurrences

We get

$$\begin{aligned}
g_k &= 3\left[g_{k-1}\right] + 2^k \\
&= 3\left[3g_{k-2} + 2^{k-1}\right] + 2^k \\
&= 3^2\left[g_{k-2}\right] + 32^{k-1} + 2^k \\
&= 3^2\left[3g_{k-3} + 2^{k-2}\right] + 32^{k-1} + 2^k \\
&= 3^3 g_{k-3} + 3^2 2^{k-2} + 32^{k-1} + 2^k \\
&= 2^k \cdot \sum_{i=0}^{k} \left(\frac{3}{2}\right)^i \\
&= 2^k \cdot \frac{(\frac{3}{2})^{k+1} - 1}{1/2} = 3^{k+1} - 2^{k+1}
\end{aligned}$$

6 Recurrences

Let $n = 2^k$:

$$\begin{aligned}
g_k &= 3^{k+1} - 2^{k+1}, \text{ hence} \\
f_n &= 3 \cdot 3^k - 2 \cdot 2^k \\
&= 3(2^{\log 3})^k - 2 \cdot 2^k \\
&= 3(2^k)^{\log 3} - 2 \cdot 2^k \\
&= 3n^{\log 3} - 2n \; .
\end{aligned}$$

6 Recurrences

Bibliography

[MS08] Kurt Mehlhorn, Peter Sanders:
Algorithms and Data Structures — The Basic Toolbox,
Springer, 2008

[CLRS90] Thomas H. Cormen, Charles E. Leiserson, Ron L. Rivest, Clifford Stein:
Introduction to algorithms (3rd ed.),
MIT Press and McGraw-Hill, 2009

[Liu85] Chung Laung Liu:
Elements of Discrete Mathematics
McGraw-Hill, 1985

The Karatsuba method can be found in [MS08] Chapter 1. Chapter 4.3 of [CLRS90] covers the "Substitution method" which roughly corresponds to "Guessing+induction". Chapters 4.4, 4.5, 4.6 of this book cover the master theorem. Methods using the characteristic polynomial and generating functions can be found in [Liu85] Chapter 10.