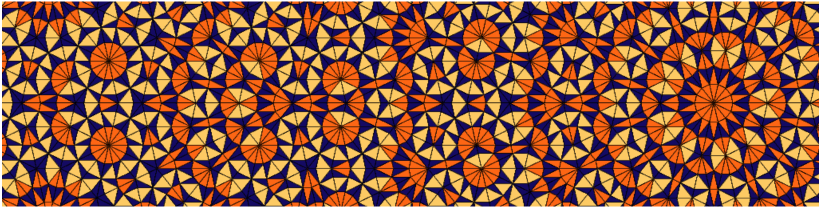
# Panorama der Mathematik und Informatik

## 7: Don Knuth I

Dirk Frettlöh
Technische Fakultät / richtig einsteigen

28.4.2015



Zwei Nachträge zu Literatur:

Warum LaTeX?

- Weil man muss (Abschlussarbeit, Fachartikel)
- Plattformunabhängig
- Zeitlos (Dokumente von vor 30 Jahren können auch in 30 weiteren Jahren noch geöffnet und bearbeitet werden)
- Top-Schriftsatz
- Freie Software

Warum nicht LaTeX?

- Weil man Word benutzen muss (Verwaltung, große Unternehmen)
- Weil man's erst lernen muss
- Weil man den Text selbst gestalten möchte
- Weil man auf drag-and-drop steht

# Zusf. Literatur in Informatik und Mathe

- Die Forschungleistung eines Wissenschaftlers können am besten Experten auf demselben Gebiet beurteilen
- Das Nächstbeste: Publikationen ansehen
  - "Gute" Zeitschrift/Tagung?
  - Viele Publikationen?
  - Viele Zitate?
- Infos dazu:
  - dblp, math scinet, zbmath.org
  - impact factor, Hirsch-Index, andere Rankings (ERA)
- Wo findet man die Artikel:
  - Homepages der Autoren
  - Preprintserver wie arxiv.org, biorXiv.org...
  - Auf den Verlagsseiten über die Unibib
- Selberlernen eher aus Lehrbüchern (Unibib...)

# Don Knuth

Donald Ervin "Don" Knuth (geb. 10.1.1938, Milwaukee, USA)

- 1956 Aufnahme des Physikstudiums am CalTech
- Bachelor und Master in Mathe 1960
- PhD (engl. für "Dr") in Mathe 1963
- Schrieb während des Studiums einen verbesserten Assembler und Compiler für den IBM 650...
- ...und ein Programm zur Verbesserung der Basketballmannschaft


IBM

Über 100 Fachartikel zu Informatik und Mathe, etliche zu anderen Themen (Sprache, Schriftsatz, Religion), viele Bücher.
Hirsch-index 28.

**Themen** insbesondere Analyse von Algorithmen, exakte Beweise für die Laufzeit von Algorithmen.

Knuth begann 1963 *The Art of Computer Programming*.

*"It was a totally new field, [...] with no real identity. And the standard of available publications was not that high. A lot of the papers coming out were quite simply wrong. [...] So one of my motivations was to put straight a story that had been very badly told."*

Zuerst geplant war ein Buch, später eine sechs-, dann eine siebenbändige Reihe.

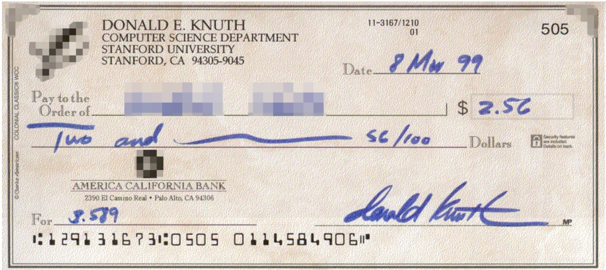
Band 1 erschien 1968, Band 2 1969, Band 3 1973.

Teil A von Band 4 erschien 2011.
Teil B teilweise als Download erhältlich.

Teil 5 ist für 2020 geplant.

Beispiele für Knuths schrägen Humor:

- In Band 1, Index:
  - “Circular definition: see definition, circular”
  - “Definition, circular: see circular definition”
- Wer einen Fehler in TAOCP findet, bekommt 2,56 US\$ ($2^{10}$ Cent = ein hexadezimaler Dollar)
- Früher als echten Scheck, heute als Scheck der “Bank von San Serriffe”
- “Beware of bugs in the above code; I have only proved it correct, not tried it.”



Knuth's Homepage: `http://cs.stanford.edu/~uno`

$\TeX$ schrieb er, weil er mit dem Schriftsatz ("Fotosatz", neue Drucktechnik) der Neuauflage von Band 2 von TAOCP unzufrieden war (1976).

*"The first goal was quality: we wanted to produce documents that were not just nice, but actually the best. [...] The second major goal was archival: to create systems that would be independent of changes in printing technology as much as possible. [...] I wanted to design something that would be still usable in 100 years."*

Knuth lernte sehr viel über Schriftsatz und -typen.
1986 war TEX (nach 9 Jahren) fertig.
Seitdem hat sich es als Standard für viele wiss. Texte durchgesetzt.

Heute benutzt man LATEX(Lamport-TEX), eine benutzerfreundlichere Version (i. Wes. Makros: documentclass, section, bibliography...).

Dazu entwickelte Knuth eine eigene *Schriftfamilie*, "Computer Modern".

Die Kunst der Textgestaltung: **Typographie**.

- Optische Wirkung: schön, passend zum Text, oder aber: auffällig, witzig...
- Lesbarkeit
- Ökonomie
- Wiedererkennungswert

*Schriftklassen*: Antiqua ("serif", also mit Serifen), Groteske ("sans serif", ohne Serifen), `Monospace`, ...

*Schnitte*: **fett**, *kursiv*, SMALL CAPITALS,...

*Schriftfamilie (typeface)*: Eine Schrift mit einigen ihrer Schnitte. (Falls viele Schnitte: auch *Schriftsippe*)

*Schriftart (font)*: Eine Schrift (z.B. Arial). Früher sogar: eine Schrift bestimmter Größe (z.B. Arial 11pt) (1 pt $= \frac{1}{72}$ inch $=$ 0,35mm)

Zum Wiedererkennungswert... welche Marke?


Design.
Limousinen


(“Corporate A”, bzw Schriftsippe “Corporate”)

Welche Marke?

**The quick brown fox jumps over the lazy dog.**

**THE QUICK BROWN FOX JUMPS ... LAZY DOG**


UND WIE
SCHLÄFST DU?
Alles für deinen guten Schlaf
Noch mehr Betten


UND WIE
SCHLÄFST DU?
Alles für deinen guten Schlaf
Noch mehr Betten
IKEA

( “Verdana” )

Welche Marke?

Frankfurter

Frankfurter

Felsquellwasser

Frankfurter

Felsquellwasser

Krombacher

(Variante von “Winston Condensed”)

Schriftarten:

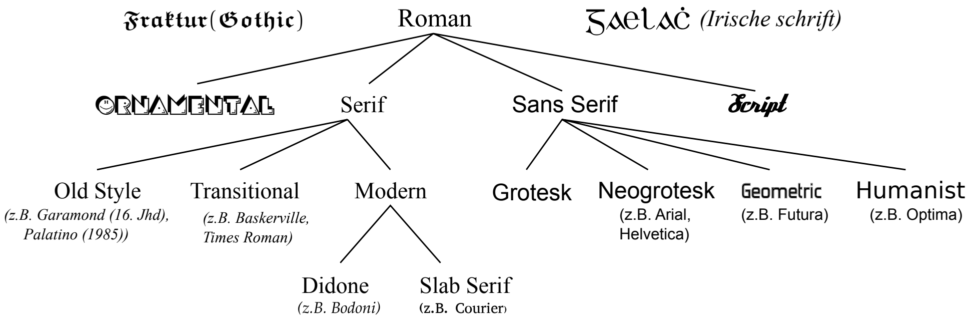
THE QUICK BROWN FOX JUMPS OVER THE LAZY DOG.
The quick brown fox jumps over the lazy dog.
THE QUICK BROWN FOX JUMPS OVER THE LAZY DOG.
The quick brown fox jumps over the lazy dog.
The quick brown fox jumps over the lazy dog.
The quick brown fox jumps over the lazy dog.
The  quick brown fox jumps over the lazy dog.
THE QUICK BROWN FOX JUMPS OVER THE LAZY DOG.
The quick brown fox jumps over the lazy dog.
THE QUICK BROWN FOX JUMPS OVER THE LAZY DOG.
The quick brown fox jumps over the lazy dog.
The quick brown fox jumps over the lazy dog.

Deutliche Unterschiede bzgl. Lesbarkeit, Platzverbrauch, Auffälligkeit...

Wie bringt man Ordnung da rein?

Die Klassifikation ist nicht ganz klar: Es gibt verschiedene Systeme. Z.B. Vox-ATypI classification, DIN 16518... hier eine grobe Einteilung:

*Lateinische Buchstaben*



Manche Namen bezeichnen einfach einzelne Schriftarten, z.B. Comic Sans MS.

Manches sagt mehr:

$$\left.\begin{array}{l}\text{Monotype}\\ \text{Linotype}\\ \ldots\end{array}\right\}\ \text{Times (New) Roman}\ \left\{\begin{array}{l}\mathrm{roman}\\ \textit{Italic}\\ \textbf{Bold}\\ \ldots\end{array}\right.$$

Computer Modern: von Don Knuth.
Linotype: Firma.

$$\left.\begin{array}{l}\text{ITC}\\ \text{Monotype}\\ \text{Adobe}\\ \ldots\end{array}\right\}\ \text{Garamond}\ \left\{\begin{array}{l}\mathrm{Roman}\\ \textit{Italic}\\ \textbf{Bold}\\ \ldots\end{array}\right.\qquad \text{(evtl. "condensed", "light")}$$

ITC, Monotype, Adobe: Firmen.
*Condensed*: Schmaler; *light*: dünner bzw "grauer".

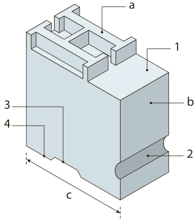
Die Schriftart Garamond (von Claude Garamond, 1499-1561) ist ein vielfach kopierter Klassiker. Fast alle Bücher und Zeitungen sind damit oder mit einem sehr ähnlichen Typ (Times, Adobe Garamond, ITC Garamond,...) gesetzt. Besonderheiten:

g W (Untere Schleife des g! W als zwei überlagerte V)

Relativ kleine Schleifen im a e

Don Knuth entwickelte mit TEX zusammen auch ein Programm METAFONT (zur digitalen Beschreibung von Schriftarten) und eine Schriftfamilie: Computer Modern.

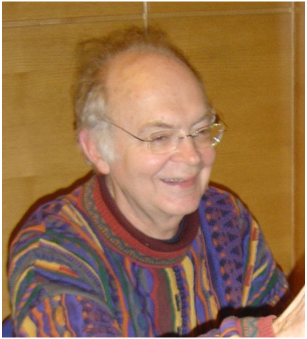
Ganz früher: Bleilettern. Jede Schriftart und -größe hat einen Satz von “Master”-buchstaben, Stempel aus Stahl (“Patrize”). Patrize schlägt buchstabenförmige Vertiefung in “Matrize” (Kupfer), diese dient als Gussform für “Letter”.




Mit dem Aufkommen des Digitaldrucks konnte man (theoretisch) die Schriftart für eine Größe vorgeben, das skalieren sollte dann automatisch gehen. Das ist aber tricksig: einfach verkleinern liefert schlechte Ergebnisse, Strichdicken und Serifenlängen und -breiten müssen angepasst werden.

Das macht Don Knuths METAFONT. (zeigen: D. Knuth: “Digital Typography” und “Mathematical Typography”)

# Don Knuths Arbeiten



- Rechnerarchitektur: RISC-Befehlssatz MIX, MMIX
- Programmiersprache: WEB, CWEB
- Computerschriftsatz: $\TeX$, METAFONT, Computer modern
- Programmierung: *The Art of Computer Programming*
- Analyse von Algorithmen: Avg.-Case und Worst-Case Laufzeit
- Diskrete Mathematik: Kombinatorik, Graphentheorie, Zahlentheorie, Algebra

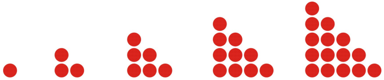
Eine Arbeit von Don Knuth soll jetzt im Detail vorgestellt und eingeordnet werden. Dazu müssen wir etwas ausholen...

**Kombinatorik:** Dinge zählen.

**Bsp.:** In wie viele Teile kann eine Pizza mit $n$ geraden Schnitten höchstens geteilt werden?

| Schnitte | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Teile | 2 | 4 | 7 | 11 | 16 | 22 | 29 | 37 | 46 | 56 | 67 | 79 | 92 | 106 |

Man kann zeigen: $n$ Schnitte, $\frac{n(n+1)}{2} + 1$ Teile $= \binom{n+1}{2} + 1$ Teile.

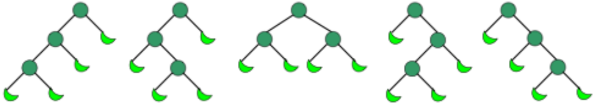
Die $\binom{n+1}{2}$ heißen **Dreieckszahlen**: 

**Bsp.:** Anzahl der Möglichkeiten, ein Produkt mit $n+1$ Faktoren (sinnvoll !) zu klammern:

$$2 \cdot (3 \cdot ((4 \cdot 5) \cdot 6)), \quad 2 \cdot (3 \cdot (4 \cdot (5 \cdot 6)), \quad (2 \cdot 3) \cdot ((4 \cdot 5) \cdot 6)), \ldots$$

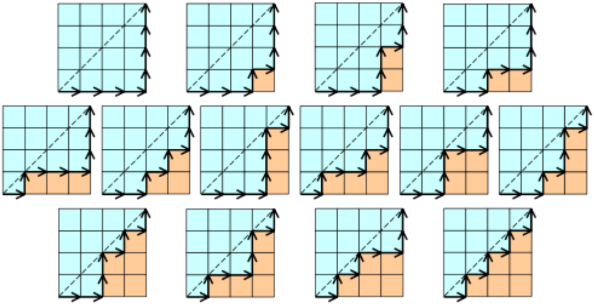
| Faktoren | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Mögl. | 1 | 2 | 5 | 14 | 42 | 132 | 429 | 1430 | 4862 | 16796 |

**Bsp.:** Anzahl der binären Wurzelbäume mit $n+1$ Blättern
(kein Knoten hat nur einen Nachfolger, rechts $\neq$ links)



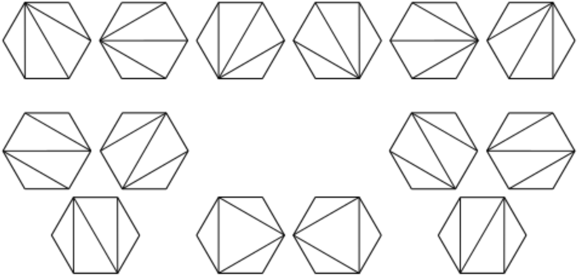
| Faktoren | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Mögl. | 1 | 2 | 5 | 14 | 42 | 132 | 429 | 1430 | 4862 | 16796 |

**Bsp.:** Anzahl der monotonen Wege in einem $n \times n$-Quadrat von unten links nach oben rechts, die nie die Diagonale überschreiten.



| Faktoren | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Mögl. | 1 | 2 | 5 | 14 | 42 | 132 | 429 | 1430 | 4862 | 16796 |

**Bsp.:** Anzahl der Möglichkeiten, ein $n + 2$-Eck mit geraden Schnitten (nur von Ecke zu Ecke) in Dreiecke zu zerteilen.



| Faktoren | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Mögl. | 1 | 2 | 5 | 14 | 42 | 132 | 429 | 1430 | 4862 | 16796 |

Die Zahlen aus den letzten vier Beispielen heißen **Catalan-Zahlen**.

Exercise 6.19 in

**Richard P. Stanley:** *Enumerative combinatorics* Band 2, Cambridge University Press 1999

listet 66 Zählprobleme auf, die als Lösung die Catalanzahlen haben.

(Im Netz: Fortsetzung, 141 weitere Zählprobleme, deren Lösung die Catalanzahlen sind)

# Online Encyclopedia of Integer Sequences

`oeis.org`

Zeigen:

- Fibonaccizahlen
- Catalanzahlen
- Kolakoski
- 1,2,3,4,5,6,7,8,9
- 1,1,4,7,19,40,97,217,508

April 2015: ungefähr 250 000 Einträge

Bei solchen **kombinatorischen** Problemen sucht man:

- Rekursionsgleichung (gut)
- Erzeugende Funktion (besser)
- Geschlossene Formel (am besten)

**Catalanzahlen:** Rekursion:

$$C_n = C_0 C_{n-1} + C_1 C_{n-2} + \cdots + C_{n-1} C_0.$$

**Erzeugende Funktion:**

$$c(x) = \frac{2}{1 + \sqrt{1-4x}}$$

**Geschlossene Formel:**[1]

$$C_n = \frac{1}{n+1}\binom{2n}{n} = \frac{(2n)!}{(n+1)!\,n!} = \prod_{k=2}^{n} \frac{n+k}{k}$$

[1] etwas, wo nur $n$ eingesetzt werden muss

# Erzeugende Funktionen

Wie findet man das? Ein Weg: Erzeugende Funktionen!

Dazu zunächst Wiederholung Mathe 1/2: Ein zentrales Thema:

**(Viele) Funktionen lassen sich als Potenzreihen schreiben.**

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

- $\frac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 + \cdots = \sum\limits_{n=0}^{\infty} x^n \quad (|x| < 1)$
- $e^x = \sum\limits_{n=0}^{\infty} \frac{1}{n!} x^n \quad (x \in \mathbb{R})$
- $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \quad (x \in \mathbb{R})$
- $\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!} \quad (x \in \mathbb{R})$
- $x^2 = 0 + 0 \cdot x + 1 \cdot x^2 + 0 \cdot x^3 + 0 \cdot x^4 + \cdots$

Wichtig in dem Zusammenhang: **Taylorreihe**.

Ist $f$ unendlich oft differenzierbar, so ist die Taylorentwicklung (um den Punkt $a$)

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots$$

$$\text{bzw. } f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

$$\text{also für } a = 0: \; f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}x^n$$

**Bsp:**

- $e^x$
- $x^2$ (!)

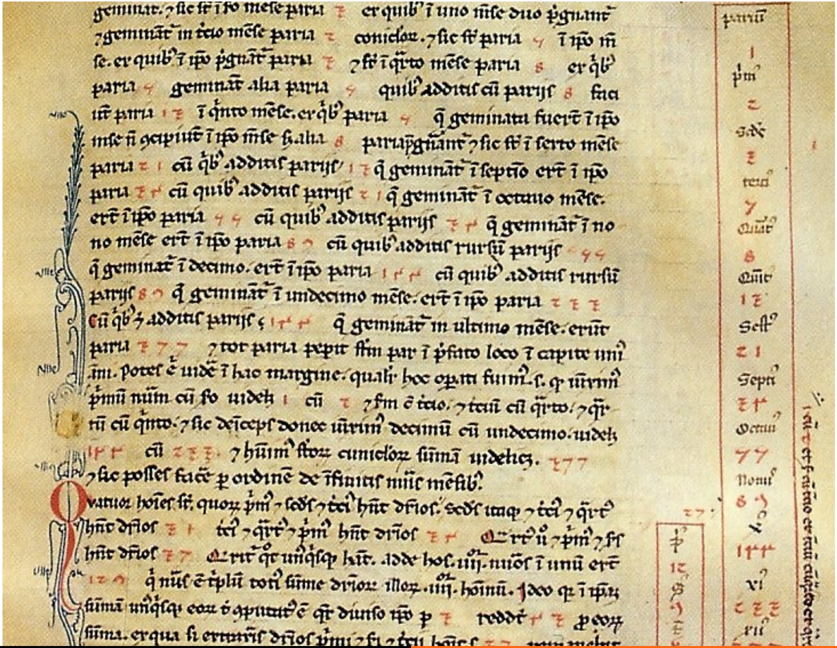
Bei kombinatorischen Problemen mit Zählwerten $a_0, a_1, a_2, \ldots$ sucht man nun $f(x)$ mit $f(x) = \sum_{n=0}^{\infty} a_n x^n$.

Erklärung am Beispiel der **Fibonaccizahlen**

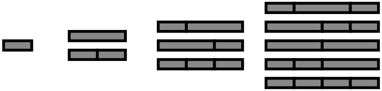
**Def.:** $f_{n+1} = f_n + f_{n-1}, \quad f_0 = 1,\ f_1 = 1$

1 junges Hasenpaar in Jahr 0 wird
1 altes Hasenpaar in Jahr 1 wird
1 altes und 1 junges Hasenpaar in Jahr 2 wird
2 alte und 1 junges Hasenpaar in Jahr 3 wird
3 alte und 2 junge Hasenpaare in Jahr 4 wird
5 alte und 3 junge Hasenpaare in Jahr 5...

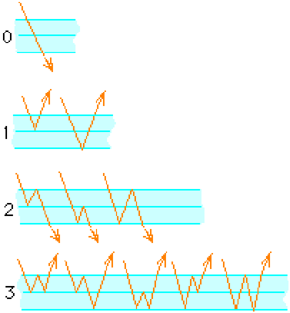
Insgesamt: 1,1,2,3,5,8,... Paare.



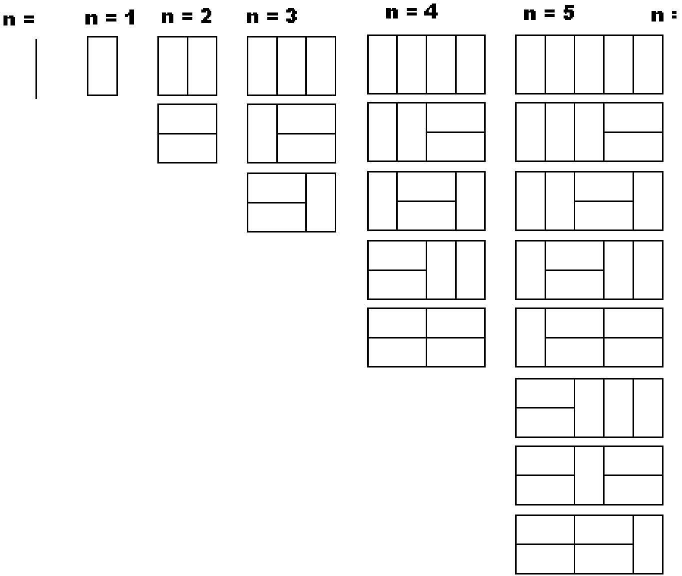
Oder: Anzahl der Möglichkeiten, ein Intervall der Länge $n$ in Intervalle der Länge 1 und 2 zu teilen:



Oder Zahl der Möglichkeiten, wie ein Lichtstrahl in einem Doppelglasfenster $(n+1)$-mal reflektiert werden kann $(\rightarrow)$



...oder: Anzahl der Möglichkeiten, mit $2 \times 1$-Dominos ein $2 \times n$-Rechteck zu legen.



**Ansatz:** $F(x) = \sum_{n=0}^{\infty} f_n x^n$.

$$
\begin{aligned}
F(x) &= 1 + x + \sum_{n=2}^{\infty} f_n x^n && (1)\\
&= 1 + x + \sum_{n=0}^{\infty} f_{n+2} x^{n+2} && (2)\\
&= 1 + x + \sum_{n=0}^{\infty} (f_{n+1} + f_n) x^{n+2} && (3)\\
&= 1 + x + \sum_{n=0}^{\infty} f_{n+1} x^{n+2} + \sum_{n=0}^{\infty} f_n x^{n+2} && (4)\\
&= 1 + x + x \sum_{n=0}^{\infty} f_{n+1} x^{n+1} + x^2 \sum_{n=0}^{\infty} f_n x^n && (5)\\
&= 1 + x + x \sum_{n=1}^{\infty} f_n x^n + x^2 F(x) && (6)\\
&= 1 + x \sum_{n=0}^{\infty} f_n x^n + x^2 F(x) = 1 + xF(x) + x^2 F(x) && (7)
\end{aligned}
$$

Also

$$F(x)(1-x-x^2)=1, \text{ also } F(x)=\frac{1}{1-x-x^2}$$

In `Maple`: `taylor`$(1/(1-x-x^2), x=0, 8)$:

$$1+x+2x^2+3x^3+5x^4+8x^5+13x^6+21x^7+O(x^8)$$

Klappt!

$F$ ist die gesuchte **erzeugende Funktion**.

Für die **Geschlossene Formel:** $F$ vereinfachen. Z.B. zu

- $\frac{1}{1-7x}=\sum_{n=0}^{\infty}7^n x^n$ (dann wäre $f_n=7^n$), oder
- $\frac{21}{1-7x}+\frac{2}{1+5x}=3\cdot 7\sum_{n=0}^{\infty}7^n x^n+2\sum_{n=0}^{\infty}(-5)^n x^n$
  (dann wäre $f_n=3\cdot 7^{n+1}+2\cdot(-5)^n$)

(Forts. folgt)