

3 C-rekursive Folgen – Theorie

Lineare Rekursionen mit konstanten Koeffizienten, kurz C-Rekursionen, sind die einfachsten, aber zugleich auch wichtigsten Rekursionsbeziehungen überhaupt:

- *einfach*, weil sich diese Thema mit Mitteln der Linearen Algebra, also Vektorraum, Basis, lineare Transformation, Eigenwerte usw., komplett beherrscht lässt. Damit sind sie ein hervorragendes Beispiel für die allgemeine Regel:

“Wann immer Du ein Problem modellieren willst/musst mit dem Ziel *quantitativer* Aussagen, versuche es zuerst mit den Mitteln der Linearen Algebra.”

Lineare Gleichungssysteme Lösen ist eine der elementarsten Techniken, und für alles, was damit zusammenhängt, stehen riesige Arsenale mächtiger Methoden bereit.

Dies praktizieren viele Bereiche der exakten Wissenschaften mit durchschlagendem Erfolg. Um nur einige zu nennen: Physik, Nachrichtentechnik, Systemtheorie, Optimierung, stochastische Prozesse, ...

- *wichtig*, weil solche Rekursionen in verschiedensten Zusammenhängen auftreten: das reicht in der Informatik von der Behandlung von Komplexitätsproblemen für Algorithmen und Formale Sprachen bis hin zur Generierung von Pseudo-Zufallsfolgen. Es ist wichtig, dass man weiss, was zu tun ist, wenn man einer Situation begegnet, in der das relevant ist.

Die folgenden beiden Kapitel sollen die gemeinsamen Mechanismen erläutern, die hierbei am Werk sind.

3.1 Zwei Beispiele zur Motivation

3.1.1 Die Fibonacci-Rekursion

Die klassischste aller Rekursionen ist die der FIBONACCI-Zahlen F_n

$$F_0 = 0, \quad F_1 = 1, \quad F_n = F_{n-1} + F_{n-2} \quad (n \geq 2)$$

mit der Folge der ersten Werte

n	0	1	2	3	4	5	6	7	8	9	10
F_n	0	1	1	2	3	5	8	13	21	34	55

Wegen ihrer engen Verbindung zu Themen wie “goldener Schnitt” und “euklidischer Algorithmus” und “Kettenbruchentwicklungen” und “orthogonale Polynome” ist dies sicher die meiststudierte rekursive Zahlenfolge in der Geschichte der Mathematik (mit Auswirkungen in Kunst, Biologie u.v.a.m). Ihr erstes Auftreten in dem epochalen Buch *LIBER ABACI*¹ des LEONARDO VON PISA alias FILIUS BONACCI um 1200 manifestiert sich als älteste überlieferte Erwähnung einer solchen rekursiven Zahlenfolge überhaupt. Für die Liebhaber dieser und ähnlicher Folgen gibt es sogar eine eigene Zeitschrift, das *FIBONACCI QUARTERLY* und spezielle Tagungen zu diesem Thema.

Als *Goldenen Schnitt* bezeichnet man das Verhältnis zweier Strecken der Länge a und b , wenn – es sei $a > b$ angenommen, das Verhältnis der Summe $a + b$ zur grösseren Strecke a das gleiche ist wie das der längeren Strecke a zur kürzeren Strecke b , also

$$\frac{a+b}{a} = \frac{a}{b}.$$

Gleichwertig dazu ist, dass der Bruch $\frac{a}{b}$ eine Lösung der quadratischen Gleichung

$$X^2 = X + 1$$

ist. Die beiden Lösungen sind

$$\phi = \frac{1 + \sqrt{5}}{2} = 1.6180339887\dots, \quad \hat{\phi} = \frac{1 - \sqrt{5}}{2} = -0.6180339887\dots$$

und die Lösung ϕ ist der Goldene Schnitt. Man findet leicht durch Induktion heraus, dass zwischen den Fibonacci-Zahlen F_n und dem Goldenen Schnitt ϕ folgender Zusammenhang besteht

$$F_n = \frac{\phi^n - \hat{\phi}^n}{\sqrt{5}}.$$

wegen $|\phi| > 1$ und $|\hat{\phi}| < 1$ zeigt dies, dass die Folge $(F_n)_{n \geq 0}$ *exponentiell schnell wächst*, eben wie ϕ^n .

Einen anderen interessanten Zusammenhang erkennt man, wenn man die sukzessiven Quotienten F_{n+1}/F_n betrachtet und der Grösse nach anordnet:

$$\frac{1}{1} < \frac{3}{2} < \frac{8}{5} < \frac{21}{13} < \dots < \frac{34}{21} < \frac{13}{8} < \frac{5}{3} < \frac{2}{1}$$

¹Dieses Buch ist wegen der Zahlenfolge ganz interessant, aber die wahre Bedeutung geht sehr viel weiter: es ist eines der ersten und vielleicht das wichtigste Zeugnis für das Aufkommen der indisch-arabischen Mathematik in Europa um 1200, mit dezimalem Zahlensystem und Null und bis heute praktizierten Rechenverfahren – den “Algorithmen”, genannt nach AL KHWARIZM, einem Mathematiker an der *Schule der Weisheit* in Bagdad um das Jahr 800. Die Tatsache, dass auch das Wort *Algebra*, aus dem Titel eines Buches von AL KHWARIZM entnommen, auf diesem Weg zu uns gekommen ist, spricht für die epochale Bedeutung dieses Vorgangs.

Diese Verhältnisse ergeben sich aus der Formel

$$F_{n+1} F_{n-1} - F_n^2 = (-1)^n,$$

die man ebenfalls leicht per Induktion beweisen kann. Sie zeigt letztlich, dass

$$\lim_{n \rightarrow \infty} \frac{F_{n+1}}{F_n} = \phi$$

gilt, wobei der Grenzwert von rechts und links “eingeschachtelt” wird.

3.1.2 Das Frisbee-Problem

Das Frisbee-Problem lässt sich als ein Transitionssystem modellieren, bei dem es drei Zustände gibt, die besagen, welches zu einem gegebenen Zeitpunkt der (kürzeste von zwei Möglichkeiten) Abstand der beiden Frisbee-Scheiben in einem Graphen von 5 Knoten auf einem Kreis ist. Also

“1” : die Scheiben haben Abstand 1 (= Startzustand)

“2” : die Scheiben haben Abstand 2

“3” : die Scheiben haben Abstand 0 (= Endzustand)

Zählt man die Anzahl $a_{i,j}$ der Möglichkeiten, von einem Zustand “ i ” in einen Zustand j zu gelangen, so ergibt das in Matrixdarstellung

$$A = [a_{i,j}]_{1 \leq i,j \leq 3} = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

Wenn man wissen will, wie viele Möglichkeiten es gibt, in n Spielzügen von einem Zustand “ i ” in einen Zustand j zu gelangen, so findet man diese Information in der n -ten Potenz der Matrix A , also

$$A^n = [a_{i,j}^{(n)}].$$

Beispielsweise ist

$$A^2 = \begin{bmatrix} 10 & 5 & 1 \\ 5 & 5 & 3 \\ 0 & 0 & 1 \end{bmatrix} \quad A^3 = \begin{bmatrix} 35 & 20 & 6 \\ 20 & 15 & 8 \\ 0 & 0 & 1 \end{bmatrix} \quad A^4 = \begin{bmatrix} 125 & 75 & 26 \\ 75 & 50 & 23 \\ 0 & 0 & 1 \end{bmatrix}$$

$$A^{10} = \begin{bmatrix} 278125 & 171875 & 65626 \\ 171875 & 106250 & 40623 \\ 0 & 0 & 1 \end{bmatrix}$$

$$A^{20} = \begin{bmatrix} 106894531250 & 66064453125 & 25234375001 \\ 66064453125 & 40830078125 & 15595703123 \\ 0 & 0 & 1 \end{bmatrix}$$

Insbesondere interessieren diese Anzahlen für den Ausgangspunkt “ i ”=“1”. Das ist jeweils die erste Zeile der Matrix A^n . Ordnet man für $n = 0, 1, 2, 3, \dots$ diese Zahlen spaltenweise an, so erhält man das Schema

$$\mathbf{h} = \left[h_i^{(n)} \right]_{1 \leq i \leq 3, n \geq 0} = \begin{bmatrix} 1 & 3 & 10 & 35 & 125 & 450 & 1625 & 5875 & 21250 & \dots \\ & 1 & 5 & 20 & 75 & 275 & 1000 & 3625 & 13125 & \dots \\ & & 1 & 6 & 26 & 101 & 376 & 1376 & 5001 & \dots \end{bmatrix}$$

Das erhält man auch durch Ausfüllen des Trellis-Diagramms, siehe Abbildung 1. Das ist nichts anderes als ein Schema zur iterativen Berechnung von

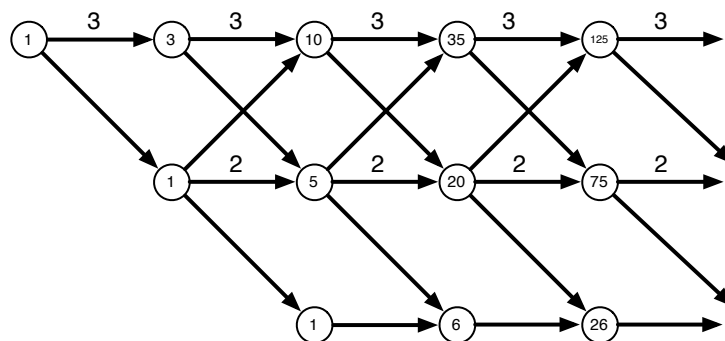


Abbildung 1: Trellis-Diagramm für das Frisbee-Spiel

$$\mathbf{h} = \left[h_i^{(n)} \right]_{1 \leq i \leq 3, n \geq 0} = ([1 \ 0 \ 0] \cdot A^n)^\top \quad \text{für } n = 0, 1, 2, 3, \dots$$

Von Interesse ist nun die Frage, wie sich die beiden Folgen

$$\mathbf{h}_1 = \left(h_1^{(n)} \right)_{n \geq 0} = [1, 3, 10, 35, 125, 450, 1625, \dots]$$

$$\mathbf{h}_2 = \left(h_2^{(n)} \right)_{n \geq 0} = [0, 1, 5, 20, 75, 275, 1000, \dots]$$

für $n \rightarrow \infty$ verhalten.

Auf experimenteller Ebene kann man feststellen, dass die (ersten paar Werte der) beiden Folgen derselben Rekursion genügen:

$$h_i^{(n)} = 5 \cdot h_i^{(n-1)} - 5 \cdot h_i^{(n-2)} \quad (i = 1, 2, n \geq 2)$$

und sich durch die Startwerte unterscheiden:

$$h_1^{(0)} = 1, h_1^{(1)} = 3, \quad h_2^{(0)} = 0, h_2^{(1)} = 1.$$

Die dritte Folge

$$\mathbf{h}_3 = \left(h_3^{(n)} \right)_{n \geq 0} = [0, 0, 1, 6, 26, 101, 376, 1376, \dots]$$

genügt einer etwas anderen (“inhomogenen”) Rekursion

$$h_3^{(n)} = 5 \cdot h_3^{(n-1)} - 5 \cdot h_3^{(n-2)} + 1 \quad (n \geq 2)$$

mit den Startwerten

$$h_3^{(0)} = 1, h_3^{(1)} = 0.$$

Das liegt an der Sonderrolle als Zustand, von dem aus man nicht zu den anderen zurückkommen kann. Man kann auch eine “homogene” Rekursion für die Folge \mathbf{h}_3 finden:

$$h_3^{(n)} = 6 \cdot h_3^{(n-1)} - 10 \cdot h_3^{(n-2)} + 5 \cdot h_3^{(n-3)} \quad (n \geq 3)$$

mit den Startwerten

$$h_3^{(0)} = 0, h_3^{(1)} = 0, h_3^{(2)} = 1.$$

Wie kommt es zu diesen Rekursionen? Als Hinweis: man sollte sich das *charakteristische Polynom* der Matrix A anschauen:

$$\chi_A(z) = \det(z \cdot \mathbb{I}_3 - A) = z^3 - 6z^2 + 10z - 5 = (z - 1) \cdot (z^2 - 5z + 5)$$

Diese Polynome haben genau die Koeffizienten, wie sie in der Rekursionen auftreten. Dabei wird das Geschehen auf den beiden Zuständen “1” und “2” durch das Polynom $z^2 - 5z + 5$ “beschrieben”. Der “Sonderstatus” des Zustandes “3” macht sich durch den zusätzlichen Faktor $z - 1$ bemerkbar.

3.2 Zwei Vorbemerkungen

3.2.1 Lineare Rekursionen 1. Ordnung

Ist eine Folge $\mathbf{x} = (x_n)_{n \geq 0}$ von komplexen Zahlen gegeben durch die lineare Rekursion 1. Ordnung

$$x_{n+1} = a \cdot x_n \quad (n \geq 0) \quad \text{mit Anfangswert } x_0,$$

mit einer komplexen Konstanten a , so gilt offensichtlich

$$x_n = a^n \cdot x_0 \quad (n \geq 0).$$

Insbesondere gilt dann auch

$$|x_n| = |a|^n \cdot |x_0| \quad (n \geq 0).$$

Über das asymptotische Verhalten von x hat man daher einen einfachen Überblick

- Falls $x_0 = 0$ ist, ist auch $x_n = 0$ für alle $n \geq 0$.
- Falls $x_0 \neq 0$ ist:
 - Falls $|a| > 1$ ist, gilt $|x_n| \uparrow \infty$; die Folge wächst exponentiell schnell.
 - Falls $|a| = 1$ ist, gilt $|x_n| = |x_0|$ für alle $n \geq 0$; die Folge bleibt beschränkt.
 - Falls $|a| < 1$ ist, gilt $|x_n| \downarrow 0$; die Folge schrumpft exponentiell schnell gegen 0.

Dieses typische Verhalten wird sich auch in der allgemeinen Situation wiederfinden. Lösungen setzen sich aus diesen drei Lösungstypen additiv zusammen.

Man kann noch bemerken, dass hier a Nullstelle der “charakteristischen Gleichung” $z - a = 0$ ist.

3.2.2 Potenzen von Matrizen

Hat man eine $(k \times k)$ -Matrix A und steht vor der Aufgabe, die Potenz A^n für ein grosses n zu berechnen, so kann man - statt dies auf dem üblichen Weg (vgl. späteren Abschnitt) auszurechnen, auf folgendermassen vorgehen:

- (1) Diagonalisiere A , d.h. finde (falls das möglich ist) eine nichtsinguläre Transformation T mit

$$T \cdot A \cdot T^{-1} = D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_k \end{bmatrix}$$

wobei die $\lambda_1, \lambda_2, \dots, \lambda_k$ die Eigenwerte von A sind.

(2) Berechne

$$D^n = \begin{bmatrix} \lambda_1^n & 0 & \dots & 0 \\ 0 & \lambda_2^n & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_k^n \end{bmatrix}$$

(3) Führe die Rücktransformation aus:

$$A^n = (T^{-1} \cdot D \cdot T)^n = T^{-1} \cdot D^n \cdot T$$

Der entscheidende Vorteil liegt in dem Schritt (2): Potenzen von Diagonalmatrizen sind offensichtlich viel “billiger” zu berechnen als Potenzen von allgemeinen Matrizen. Die Kosten für das Auffinden der Matrizen der Matrix T und die Berechnung von deren Inversen – falls die überhaupt existieren, was ja nicht garantiert ist – müsste man natürlich fairerweise mit in die Kostenkalkulation einbeziehen.

3.3 Definitionen, grundlegende Eigenschaften

Definition 1. Es bezeichne $\mathbb{C}^{\mathbb{N}}$ die Menge aller unendlichen Folgen

$$\mathbf{x} = (x_0, x_1, x_2, \dots) = (x_n)_{n \geq 0}$$

von komplexen Zahlen, die, versehen mit der komponentenweisen Addition

$$\mathbf{x} + \mathbf{y} = (x_0 + y_0, x_1 + y_1, x_2 + y_2, \dots) = (x_n + y_n)_{n \geq 0}$$

und der Skalarmultiplikation mit $\alpha \in \mathbb{C}$

$$\alpha \mathbf{x} = (\alpha x_0, \alpha x_1, \alpha x_2, \dots) = (\alpha x_n)_{n \geq 0}$$

ein (unendlich-dimensionaler) komplexer Vektorraum ist. Das neutrale Element ist die Nullfolge

$$\mathbf{0} = (0, 0, 0, \dots) = (0)_{n \geq 0}.$$

Uns interessieren in der Folge spezielle endlich-dimensionale Teilräume von $\mathbb{C}^{\mathbb{N}}$.

Definition 2. Sind $\mathbf{a} = (a_1, a_2, \dots, a_k) \in \mathbb{C}^k$ mit $a_k \neq 0$, so ist durch

$$(*_n) \quad x_n = a_1 x_{n-1} + a_2 x_{n-2} + \dots + a_k x_{n-k}$$

eine *lineare Rekursion k -ter Ordnung mit konstanten Koeffizienten*, kurz: eine *C-Rekursion der Ordnung k* definiert.

Eine Folge $\mathbf{x} = (x_0, x_1, x_2, \dots) = (x_n)_{n \geq 0} \in \mathbb{C}^{\mathbb{N}}$ genügt der durch \mathbf{a} definierten Rekursion, wenn $(*_n)$ für alle $n \geq k$ gilt.

Die Menge aller bezüglich \mathbf{a} rekursiven Folgen wird mit

$$\mathcal{V}_{\mathbf{a}} = \{ \mathbf{x} \in \mathbb{C}^{\mathbb{N}}; \mathbf{x} \text{ erfüllt } (*_n) \text{ für alle } n \geq k \}$$

bezeichnet.

Das Polynom vom Grad k

$$a(z) = 1 - a_1 z - a_2 z^2 - \dots - a_k z^k$$

wird als *Rekursionspolynom* (der durch \mathbf{a} gegebenen Rekursion) bezeichnet, das Polynom (ebenfalls vom Grad k)

$$\chi_{\mathbf{a}}(z) = z^k - a_1 z^{k-1} - a_2 z^{k-2} - \dots - a_k$$

als das *charakteristische Polynom* dieser Rekursion.

Beispiel 1. Die FIBONACCI-Rekursion $F_n = F_{n-1} + F_{n-2}$ hat das Rekursionspolynom $1 - z - z^2$ und das charakteristische Polynom $z^2 - z - 1$.

Die FRISBEE-Rekursion $G_n = 5G_{n-1} - 5G_{n-2}$ hat das Rekursionspolynom $1 - 5z + 5z^2$ und das charakteristische Polynom $z^2 - 5z + 5$.

Bemerkung 1. 1. Ist $\mathbf{x} = (x_n)_{n \geq 0}$ eine \mathbb{C} -rekursive Folge mit einer durch \mathbf{a} gegebenen Rekursion der Ordnung k , so ist \mathbf{x} durch die Angabe der k "Startwerte" x_0, x_1, \dots, x_{k-1} eindeutig bestimmt.

2. Offensichtlich kann $\xi = 0$ weder eine Nullstelle des Rekursionspolynoms $a(z)$, noch eine Nullstelle des charakteristischen Polynoms $\chi_{\mathbf{a}}(z)$ sein, da der jeweilige konstante Term $\neq 0$ ist. Ansonsten gilt für $\xi \in \mathbb{C}$:

$$a(\xi) = 0 \Leftrightarrow \chi_{\mathbf{a}}(\xi^{-1}) = 0,$$

d.h. die Nullstellen von $\chi_{\mathbf{a}}(z)$ sind genau die Reziproken der Nullstellen der Nullstellen von $a(z)$.

Das überträgt sich auch auf mehrfache Nullstellen: $a(z)$ hat genau dann $\xi \in \mathbb{C}$ als t -fache Nullstelle, wenn ξ^{-1} eine t -fache Nullstelle von $\chi_{\mathbf{a}}(z)$ ist.

Satz 1. Für $\mathbf{a} = (a_1, a_2, \dots, a_k) \in \mathbb{C}^k$ mit $a_k \neq 0$ ist $\mathcal{V}_{\mathbf{a}}$ ein Untervektorraum der Dimension k von $\mathbb{C}^{\mathbb{N}}$.

Zum Beweis: Die Rekursionsbedingungen $(*_n)$ sind *lineare* Bedingungen, deshalb ist \mathcal{V}_a gegenüber Addition und Skalarmultiplikation abgeschlossen. Ein Element $\mathbf{x} = (x_n)_{n \geq 0}$ ist durch seine k Startwerte x_0, x_1, \dots, x_{k-1} eindeutig festgelegt, deshalb ist $\dim \mathcal{V}_a = k$.

Eine Basis von \mathcal{V}_a bilden beispielsweise die k Folgen $\mathbf{e}^{(j)} = \left(e_n^{(j)} \right)_{n \geq 0}$, $(0 \leq j < k)$, die durch die Werte

$$e_n^{(j)} = \delta_{j,n}$$

gegeben sind (Standardbasis). □

Satz 2. Für $\mathbf{a} = (a_1, a_2, \dots, a_k) \in \mathbb{C}^k$ mit $a_k \neq 0$ habe das charakteristische Polynom $\chi_a(z)$ k verschiedene komplexe Nullstellen $\lambda_1, \lambda_2, \dots, \lambda_k$. Dann bilden die k Folgen

$$\boldsymbol{\lambda}_j = (1, \lambda_j, \lambda_j^2, \lambda_j^3, \dots) = (\lambda_j^n)_{n \geq 0} \quad (1 \leq j \leq k)$$

eine Basis von \mathcal{V}_a .

Zum Beweis: Die Aussage $\chi_a(\xi) = 0$ schreibt sich

$$\xi^k - a_1 \xi^{k-1} - a_2 \xi^{k-2} - \dots - a_k = 0$$

und das ist (wegen $\xi \neq 0$) äquivalent zu

$$\xi^n = a_1 \xi^{n-1} + a_2 \xi^{n-2} + \dots + a_k \xi^{n-k}$$

für alle $n \geq k$. Somit sind alle angegebenen Folgen $\boldsymbol{\lambda}_j$ Elemente von \mathcal{V}_a .

Dass diese k Folgen linear-unabhängig sind, folgt daraus, dass die $(k \times k)$ -Matrix der Startwerte

$$[\lambda_j^{i-1}]_{1 \leq i, j \leq k} = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \dots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^{k-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_k & \lambda_k^2 & \dots & \lambda_k^{k-1} \end{bmatrix}$$

ein VANDERMONDE-Matrix ist, dass also

$$\det [\lambda_j^{i-1}]_{1 \leq i, j \leq k} = \prod_{1 \leq i < j \leq k} (\lambda_j - \lambda_i) \neq 0$$

gilt. □

Bemerkung 2. Der Beweis macht entscheidenden Gebrauch von der Voraussetzung, dass die Nullstellen $\lambda_1, \lambda_2, \dots, \lambda_k$ paarweise verschieden sind! Eine allgemeinere Formulierung, die auch den Fall mehrfacher Nullstellen einschliesst, wird weiter unten gegeben.

Für den Rest dieses Abschnitts wird angenommen, dass k paarweise verschiedene Nullstellen vorliegen.

Folgerung 3. Zu jeder Folge $\mathbf{x} = (x_n)_{n \geq 0} \in \mathcal{V}_{\mathbf{a}}$ gibt es eindeutig bestimmte Konstante $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathcal{C}$ mit

$$\mathbf{x} = \alpha_1 \boldsymbol{\lambda}_1 + \alpha_2 \boldsymbol{\lambda}_2 + \dots + \alpha_k \boldsymbol{\lambda}_k$$

also explizit ausgeschrieben

$$x_n = \alpha_1 \lambda_1^n + \alpha_2 \lambda_2^n + \dots + \alpha_k \lambda_k^n \quad (n \geq 0).$$

Diese Koeffizienten ergeben sich als die eindeutig bestimmte Lösung des linearen Gleichungssystems

$$[\alpha_1 \quad \alpha_2 \quad \dots \quad \alpha_k] \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \dots & \lambda_1^{k-1} \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^{k-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_k & \lambda_k^2 & \dots & \lambda_k^{k-1} \end{bmatrix} = [x_0 \quad x_1 \quad \dots \quad x_{k-1}]$$

Kommentar 3. Diese Darstellung zeigt explizit, wie das Verhalten der Folgekoeffizienten x_n für wachsendes n von den Nullstellen $\lambda_1, \lambda_2, \dots, \lambda_k$ bestimmt wird:

$$x_n = \sum_{|\lambda_i| > 1} \alpha_i \lambda_i^n + \sum_{|\lambda_j| = 1} \alpha_j \lambda_j^n + \sum_{|\lambda_\ell| < 1} \alpha_\ell \lambda_\ell^n$$

Die erste Summe enthält *exponentiell wachsende* Summanden, die zweite Summe ist *beschränkt*, die dritte Summe *schrumpft exponentiell schnell* gegen 0.

Es ist üblich, die Nullstellen so zu nummerieren, dass

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_k|$$

gilt. Das wird auch hier so gehandhabt.

Beispiel 2. Die Situation $|\lambda_1| > |\lambda_j|$ ($2 \leq j \leq k$) ist besonders interessant und in Anwendungen häufig anzutreffen: man sagt dann, dass λ_1 eine *dominierende* Nullstelle sei.

In diesem Fall kann man, falls $\mathbf{x} = (x_n)_{n \geq 0} \in \mathcal{V}_a$ mit $\alpha_1 \neq 0$ ist, x_n so darstellen:

$$x_n = \lambda_1^n \underbrace{\left(\alpha_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^n + \cdots + \alpha_k \left(\frac{\lambda_k}{\lambda_1} \right)^n \right)}_{(\dagger)}$$

wobei der Ausdruck (\dagger) wegen

$$\left| \frac{\lambda_j}{\lambda_1} \right| < 1 \quad \text{für } 2 \leq j \leq k$$

exponentiell (!) schnell gegen 0 konvergiert. Daher gilt in dieser Situation die asymptotische Aussage

$$x_n \sim \alpha_1 \lambda_1^n$$

3.4 Die beiden Probleme (revisited)

3.4.1 Die Fibonacci-Rekursion

Die FIBONACCI-Rekursion $F_n = F_{n-1} + F_{n-2}$ hat das Rekursionspolynom $F(z) = 1 - z - z^2$ und das charakteristische Polynom $\chi_F(z) = z^2 - z - 1$, das schon im Zusammenhang mit dem Goldenen Schnitt aufgetreten ist. $\phi = (1 + \sqrt{5})/2$ und $\hat{\phi} = (1 - \sqrt{5})/2$ sind die beiden Nullstellen:

$$\chi_F(z) = z^2 - z - 1 = (z - \phi)(z - \hat{\phi})$$

Es gilt also $\phi - \hat{\phi} = 1$ und $\phi \cdot \hat{\phi} = -1$. Die Glieder der FIBONACCI-Folge $(F_n)_{n \geq 0}$ müssen sich in der Form

$$F_n = \alpha_1 \phi^n + \alpha_2 \hat{\phi}^n$$

darstellen lassen. Für $n = 0, 1$ erhält man die Bedingungen

$$0 = F_0 = \alpha_1 + \alpha_2, \quad 1 = F_1 = \alpha_1 \phi + \alpha_2 \hat{\phi},$$

die sich ohne weiteres lösen lassen zu

$$\alpha_1 = \frac{1}{\sqrt{5}}, \quad \alpha_2 = -\frac{1}{\sqrt{5}}.$$

Damit hat man die eingangs erwähnte Darstellung

$$F_n = \frac{\phi^n - \hat{\phi}^n}{\sqrt{5}}$$

aus den "allgemeinen Prinzipien" hergeleitet.

3.4.2 Das Frisbee-Problem

Das Geschehen wird durch die Eigenwerte der Frisbee-Matrix

$$A = [a_{i,j}]_{1 \leq i,j \leq 3} = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

bestimmt. Es sind dies

$$\begin{aligned} \lambda_1 &= \frac{5 + \sqrt{5}}{2} = 3.618033988 \dots, \\ \lambda_2 &= \frac{5 - \sqrt{5}}{2} = 1.381966012 \dots, \\ \lambda_3 &= 1. \end{aligned}$$

Für das praktische Rechnen mit diesen Eigenwerten ist es nützlich, die folgenden Beziehungen zu beachten:

$$\lambda_1 + \lambda_2 = 5, \quad \lambda_1 - \lambda_2 = \sqrt{5}, \quad \lambda_1 \cdot \lambda_2 = 5.$$

Für das Geschehen auf den Zuständen “1” und “2” ist nur die Teilmatrix

$$A' = [a_{i,j}]_{1 \leq i,j \leq 2} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}$$

mit den Eigenwerten λ_1 und λ_2 zuständig. Die zugehörigen Folgen $\boldsymbol{\lambda}_1 = (\lambda_1^{(n)})_{n \geq 0}$ und $\boldsymbol{\lambda}_2 = (\lambda_2^{(n)})_{n \geq 0}$ genügen offensichtlich der Rekursion

$$\lambda_i^{(n)} = 5 \cdot \lambda_i^{(n-1)} - 5 \cdot \lambda_i^{(n-2)} \quad (i = 1, 2, n \geq 2).$$

Die beiden Folgen $\boldsymbol{h}_1 = (h_1^{(n)})_{n \geq 0}$ und $\boldsymbol{h}_2 = (h_2^{(n)})_{n \geq 0}$ müssen der gleichen Rekursion genügen, haben aber andere Startwerte. Es muss also Konstante $\alpha, \beta, \delta, \epsilon$ geben mit

$$\begin{aligned} \boldsymbol{h}_1 &= \alpha \cdot \boldsymbol{\lambda}_1 + \beta \cdot \boldsymbol{\lambda}_2 \\ \boldsymbol{h}_2 &= \delta \cdot \boldsymbol{\lambda}_1 + \epsilon \cdot \boldsymbol{\lambda}_2 \end{aligned}$$

Aus dem Vergleich der Anfangswerte ergibt sich

$$\begin{bmatrix} \alpha & \beta \\ \delta & \epsilon \end{bmatrix} \begin{bmatrix} 1 & \lambda_1 \\ 1 & \lambda_2 \end{bmatrix} = \begin{bmatrix} h_1^{(0)} & h_1^{(1)} \\ h_2^{(0)} & h_2^{(1)} \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix}$$

und somit

$$\alpha = \frac{\lambda_1}{5}, \quad \beta = \frac{\lambda_2}{5}, \quad \delta = \frac{1}{\sqrt{5}}, \quad \beta = -\frac{1}{\sqrt{5}}.$$

Es gilt also

$$h_1^{(n)} = \frac{1}{5} (\lambda_1^{n+1} + \lambda_2^{n+1}), \quad h_2^{(n)} = \frac{1}{\sqrt{5}} (\lambda_1^n - \lambda_2^n).$$

Mit diesen Informationen lässt sich nun die Frage nach der mittleren Spieldauer beantworten. Dafür betrachtet man einen Zufallsvariable X , bei das Ereignis “ $X = k$ ” bedeutet: “das Spiel ist nach genau k Würfeln beendet”. Für den Erwartungswert von X erhält man dann:

$$\begin{aligned} \mathbf{E}[X] &= \sum_{k \geq 0} k \cdot \mathbf{P}[X = k] \\ &= \sum_{k \geq 0} \sum_{n < k} \mathbf{P}[X = k] \\ &= \sum_{n \geq 0} \sum_{k > n} \mathbf{P}[X = k] \\ &= \sum_{n \geq 0} \mathbf{P}[\text{Spiel benötigt } > n \text{ Würfe}] \end{aligned}$$

Nun ist $\mathbf{P}[\text{Spiel benötigt } > n \text{ Würfe}]$ die Wahrscheinlichkeit dafür, dass sich das Spiel nach n Würfeln im Zustand “1” oder im Zustand “2” befindet – und diese Wahrscheinlichkeit ist nichts anderes als

$$\frac{1}{4^n} (h_1^{(n)} + h_2^{(n)}).$$

Damit gilt also

$$\begin{aligned} \mathbf{E}[X] &= \sum_{n \geq 0} \frac{1}{4^n} (h_1^{(n)} + h_2^{(n)}) \\ &= \sum_{n \geq 0} \frac{1}{4^n} \left[\frac{1}{5} (\lambda_1^{n+1} + \lambda_2^{n+1}) + \frac{1}{\sqrt{5}} (\lambda_1^n - \lambda_2^n) \right]. \end{aligned}$$

Der Rest ist Routine: man summiert die geometrischen Reihen oder zieht ein Computeralgebra-System wie `Maple` oder `Mathematica` zu Rate. In jedem Fall ergibt sich

$$\mathbf{E}[X] = 12.$$

3.5 Schieberegisterfolgen

In der digitalen Schaltungstechnik und Codierungstheorie spielen (vorwiegend binäre, aber das ist hier nicht der wesentliche Aspekt) C-rekursive Folgen eine wichtige Rolle, weil sie sehr einfach generiert werden können und viele nützliche Verwendungen haben. Die folgende Skizze soll illustrieren, wie sich eine C-Rekursion mit Hilfe eines “rückgekoppelten Schieberegisters” (LFSR=*linear feedback shift register*) darstellen bzw. implementieren lässt:

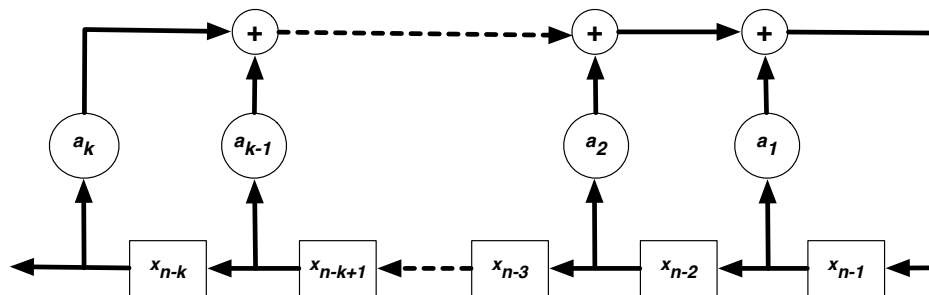


Abbildung 2: C-Rekursion und Schieberegister

Hierbei symbolisieren die Quadrate Speicherzelle, die jeweils eine Zahl aufnehmen können. Die mit a_1, \dots, a_k beschriebenen Kreise stellen Multiplikatoren dar, die übrigen, mit “+” markierten Kreise sind Addierer. Die Funktionsweise ist suggestiv:

- Zum Zeitpunkt $t = 0$ enthalten die Speicherzellen (von links nach rechts gelesen) die Startwerte $x_{k-1}, x_{k-2}, \dots, x_0$.
- Haben die Speicherzellen zu einem Zeitpunkt $t = n - 1$ die Inhalte (von links nach rechts gelesen) $x_{n-k}, x_{n-k+1}, \dots, x_{n-1}$, so werden diese in einem Takt in die jeweils linke benachbarte Zelle verschoben. Ausserdem wird der ursprüngliche Wert x_{n-k} der linkeste Zelle ausgelesen und die rechteste Zelle neu geladen mit dem Wert

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \dots + a_k x_{n-k}.$$

Zum Zeitpunkt $t = n$ sind die Zelleninhalte also $x_{n-k+1}, x_{n-k+2}, \dots, x_n$.

Arbeitet man über einem binären Alphabet, also dem zweielementigen Körper $\mathbb{F}_2 = (\mathbb{B}; \oplus, \otimes)$, wobei \oplus bzw. \otimes die Addition modulo 2 bedeuten, so ist eine derart generierte Folge notwendigerweise *periodisch*. Mit Schieberegistern aus k Speicherzellen, kann man binäre Folgen generieren, die Periodenlänge $2^k - 1$ haben: das ist

dann der Fall, wenn alle k -bit-Vektoren (ausser dem Nullvektor) zyklisch durchlaufen werden. Um zu verstehen, wie man das macht, muss man in die Polynomalgebra über endlichen Körpern einsteigen.

Folgen dieser Art wurden und werden als *Pseudozufallsfolgen* verwendet.

Betrachtet man beispielsweise das Schieberegister

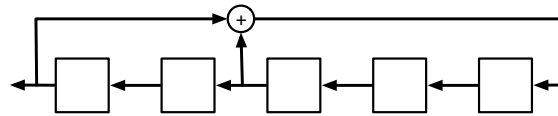


Abbildung 3: Schieberegister zur Rekursion $x_n = x_{n-3} + x_{n-5}$

so wird folgende Zustandsfolge durchlaufen (spaltenweise zu lesen):

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \quad
 \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} \quad
 \begin{bmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Hier eine Kostprobe: eine Pseudo-Zufallsfolge über \mathbb{F}_2 , die von der Rekursion $x_n = x_{n-7} + x_{n-10}$ erzeugt wurde:

0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	1	1	0	0	
1	0	0	1	1	0	1	0	0	0	0	1	0	0	1	0	1	0	1	0	0	0	0	1	1	1	1	0	1	0	1	1	
1	0	1	0	1	1	0	1	1	0	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	1	1	0	1	1	1		
0	0	1	1	0	0	0	0	1	0	1	0	1	1	0	1	0	1	1	1	0	0	0	1	1	0	1	1	1	1	1		
0	0	0	1	0	0	0	1	1	1	1	0	0	1	1	1	1	0	1	1	0	1	1	0	1	0	0	0	0	0	0		
1	0	1	0	0	0	0	1	0	1	1	0	1	0	1	0	1	0	0	0	1	1	1	1	1	0	1	1	1	1	0	0	
1	0	0	1	0	1	1	0	0	0	0	0	1	0	0	1	1	0	0	1	0	0	0	1	0	1	0	0	0	1	1	0	
1	1	0	1	1	1	0	0	0	0	0	0	1	1	1	1	0	0	0	1	1	1	0	1	1	1	1	1	1	1	0	0	
1	0	0	0	0	1	1	0	0	0	1	0	1	1	0	1	1	1	0	1	0	0	0	0	1	1	0	1	0	1	0	1	
1	0	0	1	1	1	1	0	0	1	0	1	1	0	1	1	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	1	
0	0	1	1	0	0	0	0	0	0	1	0	1	1	0	0	0	1	0	1	0	0	1	1	1	0	1	1	0	0	1	1	
1	0	0	0	1	0	1	1	1	1	1	1	0	1	0	1	0	0	0	1	0	1	1	1	0	1	1	0	1	0	1	1	
0	0	0	0	1	1	0	0	1	1	0	1	1	0	1	0	1	0	0	0	0	1	1	1	0	1	0	0	1	1	1	1	
1	0	1	0	0	1	1	0	1	0	1	0	0	1	0	0	1	1	1	0	0	0	0	0	1	1	1	1	1	0	0	1	
1	1	0	0	1	1	0	1	1	1	1	0	1	0	0	0	1	0	1	0	1	0	1	0	1	1	0	1	1	1	1	0	0
0	0	1	0	0	1	1	1	0	1	0	0	0	1	1	1	0	1	0	1	1	1	1	1	1	0	1	1	0	1	0	0	1
0	0	0	0	1	0	0	0	0	1	0	1	0	0	1	0	1	0	1	1	0	0	0	1	1	1	0	0	1	1	1	1	
1	1	1	0	1	1	0	0	0	0	1	0	0	0	1	1	0	1	0	0	1	1	1	0	0	1	0	0	1	1	1	1	
0	0	0	0	1	1	0	1	1	1	0	1	1	0	0	1	1	0	0	0	1	1	1	0	1	1	1	0	1	1	1	1	0
1	0	0	1	0	0	1	0	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	1	0	0	1	0	1	1	1	0	
1	0	0	1	0	1	1	0	1	0	0	0	1	0	0	0	1	0	1	1	0	0	1	1	0	1	0	0	1	0	1	0	
0	1	0	0	0	1	1	0	0	0	1	1	1	0	1	1	0	1	1	1	1	0	0	0	0	0	0	1	0	1	1	1	
0	0	1	0	1	0	1	1	1	0	0	1	1	1	0	1	1	1	0	1	1	1	0	0	1	1	0	0	1	1	1	0	
1	0	1	0	1	1	1	0	1	1	1	1	0	1	1	0	0	1	0	1	0	0	0	1	0	0	1	1	0	1	1	0	
0	0	1	0	0	0	0	1	1	1	0	0	1	0	1	1	1	1	0	0	1	0	1	0	0	1	1	0	0	1	1	1	
0	0	1	0	1	0	1	0	1	0	0	1	1	1	1	1	1	0	0	1	1	0	0	0	1	1	0	1	0	1	1	1	
1	0	0	1	1	0	1	0	1	1	0	1	0	0	1	1	0	0	0	1	0	0	1	0	0	1	0	1	1	0	0	0	1
0	1	1	1	1	0	1	0	1	0	1	0	1	0	1	1	1	1	1	1	1	1	0	1	0	0	0	0	0	1	0	1	
0	1	0	0	1	0	1	1	1	1	0	0	0	1	0	1	0	1	1	1	0	1	1	1	0	1	1	0	1	0	0	0	1
1	0	1	1	1	0	0	1	0	0	0	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0
0	0	1	1	1	0	0	0	0	1	1	1	1	1	1	0	1	1	1	0	0	0	1	0	0	1	1	1	1	1	0	0	0
0	1	1	0	0	1	1	1	1	1	0	1	0	1	1	0	0	1	0	1	1	0	0	1	0	0	1	0	0	1	0	0	0

3.6 Matrizen, Theorem von Cayley-Hamilton

Lineare C-Rekursionen werden durch lineare Transformationen beschrieben. Das Konzept der *Begleitmatrix* macht das. Das ist aber nur eine Umformulierung bekannter Begriffe.

Definition 3. Ist durch $\mathbf{a} = (a_1, a_2, \dots, a_k) \in \mathbb{C}^k$ mit $a_k \neq 0$, eine C-Rekursion gegeben, so bezeichnet

$$C_{\mathbf{a}} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & a_k \\ 1 & 0 & 0 & \dots & 0 & a_{k-1} \\ 0 & 1 & 0 & \dots & 0 & a_{k-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & a_2 \\ 0 & 0 & 0 & \dots & 1 & a_1 \end{bmatrix}$$

die *Begleitmatrix* (*companion matrix*) dieser Rekursion.

Satz 4. 1. Ist $\mathbf{x} = (x_n)_{n \geq 0} \in \mathcal{V}_{\mathbf{a}}$ und bezeichnet für $n \geq 0$

$$\mathbf{x}^{(n)} = [x_n \quad x_{n+1} \quad \dots \quad x_{n+k-1}]$$

den k -Vektor aufeinanderfolgender Folgeelemente, so gilt

$$\mathbf{x}^{(n+1)} = \mathbf{x}^{(n)} \cdot C_{\mathbf{a}}$$

und daher auch

$$\mathbf{x}^{(n)} = \mathbf{x}^{(0)} \cdot C_{\mathbf{a}}^n$$

2. Für das charakteristische Polynom $\chi_{\mathbf{a}}(z)$ gilt

$$\chi_{C_{\mathbf{a}}}(z) = \chi_{\mathbf{a}}(z) = \det(z \cdot \mathbb{I}_k - C_{\mathbf{a}}) = \prod_{\lambda} (z - \lambda).$$

Dabei bezeichnet \mathbb{I}_k die $(k \times k)$ -Einheitsmatrix und das Produkt läuft über alle Eigenwerte λ von $C_{\mathbf{a}}$, also die Nullstellen von $\chi_{C_{\mathbf{a}}}(z) = \chi_{\mathbf{a}}(z)$ entsprechend ihrer Vielfachheit.

3. Für das Rekursionspolynom $a(z)$ gilt

$$a(z) = \det(\mathbb{I}_k - z \cdot C_{\mathbf{a}}) = \prod_{\lambda} (1 - \lambda z).$$

Interessanter ist die Tatsache, dass und auf welchem Wege Potenzen beliebiger quadratischer Matrizen auf C-Rekursionen führen. Das ist i.w. der Inhalt eines der wichtigsten (und schönsten!) Resultate der linearen Algebra:

Theorem 5 (CAYLEY-HAMILTON).

Jede $(k \times k)$ -Matrix A erfüllt ihr charakteristisches Polynom:

$$\chi_A(A) = \mathbb{O}_k,$$

wobei \mathbb{O}_k die $(k \times k)$ -Nullmatrix ist.

Explizit gemacht:

$$A^k = a_1 A^{k-1} + a_2 A^{k-2} + \dots + a_k A^0,$$

wobei die a_1, \dots, a_k die Koeffizienten des charakteristischen Polynoms sind. Es gilt also auch – nach Multiplikation mit A^{n-k} – für jedes $n \geq k$:

$$(*_n) \quad A^n = a_1 A^{n-1} + a_2 A^{n-2} + \dots + a_k A^{n-k}$$

Schreibt man die Matrizen als

$$A^n = \left(A_{i,j}^{(n)} \right)_{1 \leq i, j \leq k}$$

und betrachtet man in der Matrixgleichung $(*_n)$ für irgendeine Position (i, j) mit $1 \leq i, j \leq k$ die Koeffizienten in der Position (i, j) , so gilt offensichtlich

$$A^n(i, j) = a_1 A^{n-1}(i, j) + a_2 A^{n-2}(i, j) + \dots + a_k A^{n-k}(i, j) \quad (n \geq k),$$

und daher:

Folgerung 6. Für jede $(k \times k)$ -Matrix A und jedes (i, j) mit $1 \leq i, j \leq k$ ist die Folge $\left(A_{i,j}^{(n)} \right)_{n \geq 0}$ C-rekursiv mit charakteristischem Polynom $\chi_A(z)$.

3.7 Reverse engineering

Bisher war die Folge $\mathbf{a} = (a_1, a_2, \dots, a_k)$ der Rekursionskoeffizienten gegeben und es wurde das Verhalten von Folgen $\mathbf{x} = (x_n)_{n \geq 0}$ untersucht, die der entsprechenden Rekursion genügen.

Interessant (und praktisch relevant in der Nachrichtentechnik, Codierungstheorie und Kryptografie) ist aber auch die umgekehrte Fragestellung:

- Gegeben eine Folge $\mathbf{x} = (x_n)_{n \geq 0}$, von der man weiss, dass sie C-rekursiv ist, aber deren Rekursionskoeffizienten man nicht kennt:
bestimme eine solche Folge $\mathbf{a} = (a_1, a_2, \dots, a_k)$ mit $\mathbf{x} \in \mathcal{V}_{\mathbf{a}}$!

Eine solche Folge muss keineswegs $\mathbf{a} = (a_1, a_2, \dots, a_k)$ sein, aber man kann zeigen, dass die *kürzeste* solche Folge, also mit minimalem k , eindeutig bestimmt ist. Für diese minimale Länge gibt es ein einfaches Kriterium und sobald diese minimale Länge k bekannt ist, kann man aus $2k$ aufeinanderfolgenden Gliedern der Folge $\mathbf{x} = (x_n)_{n \geq 0}$ auch die Koeffizienten $\mathbf{a} = (a_1, a_2, \dots, a_k)$ durch Lösung eines linearen Gleichungssystems berechnen.

Definition 4. Ist durch $\mathbf{a} = (a_1, a_2, \dots, a_k) \in \mathbb{C}^k$ mit $a_k \neq 0$, eine C-Rekursion gegeben, so bezeichnet

$$H^{(n)}(\mathbf{x}) = \begin{bmatrix} \mathbf{x}^{(n)} \\ \mathbf{x}^{(n+1)} \\ \vdots \\ \mathbf{x}^{(n+k-1)} \end{bmatrix} = \begin{bmatrix} x_n & x_{n+1} & \dots & x_{n+k-1} \\ x_{n+1} & x_{n+2} & \dots & x_{n+k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n+k-1} & x_{n+k} & \dots & x_{n+2k-2} \end{bmatrix}$$

die n -te *Hankelmatrix* der Folge \mathbf{x} .

Satz 7. *Damit ist*

$$H^{(n+1)}(\mathbf{x}) = H^{(n)}(\mathbf{x}) \cdot C_{\mathbf{a}} = \dots = H^{(0)}(\mathbf{x}) \cdot C_{\mathbf{a}}^{n+1}.$$

und

$$\det H^{(n)}(\mathbf{x}) = \det H^{(0)}(\mathbf{x}) \cdot a_k^n.$$

Satz 8. *Die durch $a(z)$ gegebene Rekursion für \mathbf{x} hat minimale Länge genau dann, wenn je k aufeinanderfolgende Vektoren $\mathbf{x}^{(n)}, \mathbf{x}^{(n+1)}, \dots, \mathbf{x}^{(n+k-1)}$ linear-unabhängig sind, d.h. wenn*

$$\det H^{(n)}(\mathbf{x}) \neq 0$$

für ein n und damit für alle $n \geq 0$.

Folgerung 9. *Beachte, dass*

$$[a_k \ a_{k-1} \ \dots \ a_1] \cdot H^{(n)}(\mathbf{x}) = \mathbf{x}^{(n+k)}$$

ist, ausgeschrieben (da $H^{(n)}(\mathbf{x})$ symmetrisch ist):

$$\begin{bmatrix} x_n & x_{n+1} & \cdots & x_{n+k-1} \\ x_{n+1} & x_{n+2} & \cdots & x_{n+k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n+k-1} & x_{n+k} & \cdots & x_{n+2k-2} \end{bmatrix} \begin{bmatrix} a_k \\ a_{k-1} \\ \cdots \\ a_1 \end{bmatrix} = \begin{bmatrix} x_{n+k} \\ x_{n+k+1} \\ \cdots \\ x_{n+2k-1} \end{bmatrix}$$

Sind die $2k$ aufeinanderfolgenden Werte $x_n, x_{n+1}, \dots, x_{n+2k-1}$ bekannt und ist weiss man, dass \mathbf{x} C-rekursiv mit der minimalen Ordnung k ist, also $\det H^{(n)} \neq 0$, so kann man daraus die a_1, a_2, \dots, a_k berechnen.

3.8 Rationale Funktionen

Es sei zunächst einmal, wie bisher,

$$a(z) = 1 - a_1 z - a_2 z^2 - \cdots - a_k z^k \quad \text{mit } a_k \neq 0$$

ein Rekursionspolynom vom Grad k , gegeben durch die Koeffizientenfolge $\mathbf{a} = (a_1, a_2, \dots, a_k)$, das eine C-Rekursion

$$f_n = a_1 f_{n-1} + a_2 f_{n-2} + \cdots + a_k f_{n-k} \quad (n \geq k)$$

der Ordnung k definiert. Jede Folge $\mathbf{f} = (f_n)_{n \geq 0}$, die dieser Rekursion genügt, also Element von $\mathcal{V}_{\mathbf{a}}$ ist, ist durch die k Anfangswerte f_0, f_1, \dots, f_{k-1} eindeutig bestimmt.

Ist nun $b(z) = b_0 + b_1 z + b_2 z^2 + \cdots + b_{k-1} z^{k-1}$ ein Polynom höchstens vom Grad $k-1$, so kann man dessen Koeffizienten b_0, b_1, \dots, b_{k-1} dazu benutzen, die k Anfangswerte f_0, f_1, \dots, f_{k-1} festzulegen:

$$\begin{aligned} f_0 &= b_0 \\ f_1 &= a_1 f_0 + b_1 \\ f_2 &= a_2 f_1 + a_2 f_0 + b_2 \\ &\vdots \\ f_{k-1} &= a_1 f_{k-2} + a_2 f_{k-3} + \cdots + a_{k-1} f_1 + b_{k-1} \end{aligned}$$

Umgekehrt sind natürlich auch die b_0, b_1, \dots, b_{k-1} durch die f_0, f_1, \dots, f_{k-1} eindeu-

tig festgelegt:

$$\begin{aligned} b_0 &= f_0 \\ b_1 &= f_1 - a_1 f_0 \\ b_2 &= f_2 - a_2 f_1 - a_2 f_0 \\ &\vdots \\ b_{k-1} &= f_{k-1} - a_1 f_{k-2} - a_2 f_{k-3} - \cdots - a_{k-1} f_1 \end{aligned}$$

Diese k Gleichungen kann man mit den Gleichungen

$$0 = f_n - a_1 f_{n-1} - a_2 f_{n-2} - \cdots - a_k f_{n-k} \quad (n \geq k)$$

elegant zusammenfassen:

$$(*) \quad b(z) = f(z) \cdot a(z),$$

wobei für eine Folge $\mathbf{f} = (f_n)_{n \geq 0}$ mit

$$f : z \mapsto f(z) = \sum_{n \geq 0} f_n z^n.$$

eine Potenzreihe (Vgl. Abschnitt 1.2.2) gemeint ist. Die Gleichung $(*)$ ist dort sinnvoll, wo diese Potenzreihe konvergiert. Der Bereich, in dem Konvergenz herrscht, ist durch den *Konvergenzradius* ρ_f gegeben (Cauchy-Hadamard-Kriterium):

$$\rho_f = \limsup_{n \rightarrow \infty} |f_n|^{1/n}.$$

Für $|z| < \rho_f$ konvergiert die Reihe, für $|z| > \rho_f$ divergiert sie. Für $|z| = \rho_f$ können beide Fälle auftreten.

Man kann zeigen, dass in dieser Situation

$$\rho_f = \min_{a(\xi)=0} |\xi|$$

gilt, und dies ist > 0 , da $a(0) = 1 \neq 0$ ist.

Bemerkung 4. Man erlaubt sich, in der eben dargestellten Situation kurzerhand

$$f(z) = \frac{b(z)}{a(z)}$$

zu schreiben und meint damit folgendes: $f(z)$ ist eine Potenzreihe, deren Koeffizienten durch die Taylorentwicklung des Quotienten $b(z)/a(z)$ um $z = 0$ gegeben

sind. Diese Taylorreihe konvergiert in jedem Kreis um den Nullpunkt, der keine Nullstelle von $a(z)$ enthält.

In Analogie zu dem Begriff *rationale Zahl* definiert man den Begriff *rationale Funktionals* (Äquivalenzklassen von) Quotienten von Polynomen.

Definition 5. 1. Auf der Menge $\langle a(X), b(X) \rangle$ aller Paare \langle Nenner, Zähler \rangle von Polynomen $a(X), b(X) \in \mathbb{C}[X]$, wobei $a(X)$ nicht das Nullpolynom sein soll, wird eine Äquivalenzrelation definiert

$$\langle a(z), b(z) \rangle \equiv \langle c(z), d(z) \rangle \Leftrightarrow a(X) \cdot d(z) = b(X) \cdot c(X)$$

Eine Äquivalenzklasse von \equiv wird als *rationale Funktion* bezeichnet.

2. In jeder Äquivalenzklasse von \equiv gibt es einen Repräsentanten $\langle a(X), b(X) \rangle$, bei den $a(X)$ und $b(X)$ keine gemeinsame Nullstelle haben, d.h. keinen gemeinsamen Teiler vom Grad ≥ 1 . Verlangt man noch, dass der niedrigste nichtverschwindende Koeffizient von $a(X)$ gleich 1 ist, so ist $f(X) = \langle a(X), b(X) \rangle$ eindeutig bestimmt: das ist die *normierte Darstellung* der zugehörigen Äquivalenzklasse.
3. Ist $f(X) = \langle a(X), b(X) \rangle$ normiert, so ist durch

$$f : \mathbb{C} \rightarrow \mathbb{C} \cup \{\infty\} : z \mapsto f(z) = \frac{b(z)}{a(z)}$$

eine Abbildung definiert, die genau für die Nullstellen von $a(X)$ den Wert ∞ annimmt. Diese Abbildung identifiziert man mit der rationalen Funktion.

4. Die Menge der rationalen Funktionen bildet unter den üblichen Regeln der "Bruchrechnung" einen Körper, den *Körper der rationalen Funktionen*, der üblicherweise mit $\mathbb{C}(X)$ bezeichnet wird.
5. Eine *strikte* rationale Funktion ist gegeben durch eine normierte Darstellung $\langle a(X), b(X) \rangle$, wobei $a(1) = 1$ ist, d.h. der konstante Koeffizient von $a(X)$ ist $= 1$. Insbesondere ist $z = 0$ also keine Nullstelle von $a(X)$. Die Menge der strikten rationalen Funktionen ist unter Summe und Produkt abgeschlossen (nicht jedoch unter Division!).

Der folgende Satz greift nur noch einmal vorher gemachte Bemerkungen auf:

Satz 10. *Stellt $f(X) = \langle a(X), b(X) \rangle$ eine strikte rationale Funktion dar und ist $\rho_f = \min\{\lambda; a(\lambda) = 0\} > 0$ der kleinste Betrag einer Nullstelle des Nennerpoly-*

noms $a(X)$, so ist ρ_f der Konvergenzradius der Reihenentwicklung von $f(z)$:

$$f(z) = \sum_{n \geq 0} f_n z^n = \frac{b(z)}{a(z)} \quad \text{für } |z| < \rho_f.$$

Nun kann man zusammenfassen:

Theorem 11 (Charakterisierung der C-rekursiven Folgen – einfache Nullstellen).

Für eine Folge $\mathbf{f} = (f_n)_{n \geq 0} \in \mathbb{C}^{\mathbb{N}}$ und ein (komplexes) Polynom

$$a(z) = 1 - a_1 z - a_2 z^2 - \dots - a_k z^k = \prod_{1 \leq j \leq k} (1 - \lambda_j z)$$

vom Grad k mit einfachen Nullstellen $\lambda_1^{-1}, \dots, \lambda_k^{-1}$ sind folgende Aussagen äquivalent:

1. $\mathbf{f} \in \mathcal{V}_{\mathbf{a}}$, d.h. \mathbf{f} ist eine C-rekursive Folge mit Rekursionspolynom $a(z)$, d.h.

$$f_n = a_1 f_{n-1} + a_2 f_{n-2} + \dots + a_k f_{n-k} \quad (n \geq k).$$

2. Es gibt Konstante $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{C}$ mit

$$f_n = \alpha_1 \lambda_1^n + \alpha_2 \lambda_2^n + \dots + \alpha_k \lambda_k^n \quad (n \geq 0).$$

3. $f(z) = \sum_{n \geq 0} f_n z^n$ ist eine rationale Funktion, d.h. es gibt ein (komplexes) Polynom $b(z)$ vom Grad $< k$ mit

$$f(z) = \frac{b(z)}{a(z)}.$$

Was für den kompletten Beweis noch fehlt, ist die Implikation von 3. zurück zu 2. Das ist aber einfach: Man macht in der Darstellung

$$f(z) = \frac{b(z)}{a(z)} = \frac{b(z)}{\prod_{1 \leq \ell \leq k} (1 - \lambda_\ell z)}$$

eine Partialbruchzerlegung und erhält

$$f(z) = \frac{b(z)}{a(z)} = \frac{b(z)}{\prod_{1 \leq \ell \leq k} (1 - \lambda_\ell z)} = \sum_{1 \leq \ell \leq k} \frac{\beta_\ell}{1 - \lambda_\ell z}$$

mit geeigneten Konstanten β_1, \dots, β_k .

Jetzt muss man nur noch die Quotienten in den Summen in geometrische Reihen entwickeln:

$$f(z) = \sum_{n \geq 0} f_n z^n = \sum_{1 \leq \ell \leq k} \beta_\ell \sum_{n \geq 0} \lambda_\ell^n z^n = \sum_{n \geq 0} \left(\sum_{1 \leq \ell \leq k} \beta_\ell \lambda_\ell^n \right) z^n$$

was per Koeffizientenvergleich

$$f_n = \sum_{1 \leq \ell \leq k} \beta_\ell \lambda_\ell^n$$

liefert.

3.9 Mehrfache Nullstellen

Setzt man nicht voraus, dass das Nennerpolynom $a(z)$ nur einfache Nullstellen hat, so müssen die Aussagen des vorigen Abschnitts verallgemeinert werden. Dieser Abschnitt enthält nur die relevanten Aussagen, die Beweise folgen im nächsten Abschnitt.

Ist also

$$a(z) = 1 - a_1 z - a_2 z^2 - \dots - a_k z^k,$$

ein Rekursionspolynom mit mehrfachen Nullstellen, dann hat auch das charakteristische Polynom

$$\chi_a(z) = z^k - a_1 z^{k-1} - \dots - a_k$$

mehrfache Nullstellen. Sind $\lambda_1 \dots \lambda_\ell$ die verschiedenen Nullstellen von $\chi_a(z)$ und t_j die Vielfachheit der Nullstelle λ_j , so ist

$$\chi_a(z) = \prod_{j=1}^{\ell} (z - \lambda_j)^{t_j}$$

die Faktorisierung von $\chi_a(z)$ in Linearfaktoren. In diesem Fall gehört zu einer Nullstelle λ mit der Vielfachheit t des charakteristischen Polynoms nicht nur die

Basisfolge $\lambda = (\lambda^n)_{n \geq 0}$, sondern insgesamt die t linear-unabhängigen Folgen

$$\begin{aligned} \lambda^{(0)} &= (1 \cdot \lambda^n)_{n \geq 0} &&= (\lambda^0, \lambda^1, \lambda^2, \lambda^3, \dots) \\ \lambda^{(1)} &= (n \cdot \lambda^n)_{n \geq 0} &&= (0 \lambda^0, 1 \lambda^1, 2 \lambda^2, 3 \lambda^3, \dots) \\ \lambda^{(2)} &= (n^2 \cdot \lambda^n)_{n \geq 0} &&= (0 \lambda^0, 1 \lambda^1, 4 \lambda^2, 9 \lambda^3, \dots) \\ &\vdots &&\vdots \\ \lambda^{(t-1)} &= (n^{t-1} \cdot \lambda^n)_{n \geq 0} &&= (0^{t-1} \lambda^0, 1^{t-1} \lambda^1, 2^{t-1} \lambda^2, 3^{t-1} \lambda^3, \dots) \end{aligned}$$

Satz 12. 1. Der Vektorraum \mathcal{V}_a hat als Basis die k Folgen

$$\lambda_j^{(s)} = (n^s \cdot \lambda_j^n)_{n \geq 0} \quad (1 \leq j \leq \ell, 0 \leq s < t_j)$$

2. Jede Folge $\mathbf{x} = (x_n)_{n \geq 0} \in \mathcal{V}_a$ hat eine eindeutige Darstellung

$$x_n = \sum_{j=1}^{\ell} p_j(n) \cdot \lambda_j^n \quad (n \geq 0)$$

wobei p_j ein Polynom vom Grad $< t_j$ ist ($1 \leq j \leq \ell$).

Der Beweis dieser Aussage ist recht technisch und wird im folgenden Abschnitt ausgeführt. Insgesamt ergibt sich:

Theorem 13 (Charakterisierung der C-rekursiven Folgen – mehrfache Nullstellen).

Für eine Folge $\mathbf{f} = (f_n)_{n \geq 0} \in \mathbb{C}^{\mathbb{N}}$ und ein (komplexes) Polynom

$$a(z) = 1 - a_1 z - a_2 z^2 - \dots - a_k z^k = \prod_{1 \leq j \leq \ell} (1 - \lambda_j z)^{t_j}$$

vom Grad k , wobei die $\lambda_1^{-1}, \dots, \lambda_{\ell}^{-1}$ die verschiedenen Nullstellen von $a(z)$ sind und λ_j^{-1} die Vielfachheit $t_j \geq 1$ hat, also $\sum_j t_j = k$, so sind folgende Aussagen äquivalent:

1. $\mathbf{f} \in \mathcal{V}_a$, d.h. \mathbf{f} ist eine C-rekursive Folge mit Rekursionspolynom $a(z)$, d.h.

$$f_n = a_1 f_{n-1} + a_2 f_{n-2} + \dots + a_k f_{n-k} \quad (n \geq k).$$

2. Es gibt Polynome $p_1(z), p_2(z), \dots, p_{\ell}(z) \in \mathbb{C}[z]$, wobei $p_j(z)$ einen Grad $< t_j$ hat, mit

$$f_n = p_1(n) \lambda_1^n + p_2(n) \lambda_2^n + \dots + p_{\ell}(n) \lambda_{\ell}^n \quad (n \geq 0).$$

3. $f(z) = \sum_{n \geq 0} f_n z^n$ ist eine rationale Funktion, d.h. es gibt ein (komplexes) Polynom $b(z)$ vom Grad $< k$ mit

$$f(z) = \frac{b(z)}{a(z)}.$$

Kommentar 5. Man kann die Bedingung fallenlassen, dass der Grad des Polynoms $b(z)$ kleiner sein muss als der Grad des Polynoms $a(z)$. Das führt nur zu bescheidenen Modifikationen, denn ist der Grad von $b(z)$ beliebig, so kann man mittels Division Polynome $q(z)$ und $r(z)$ eindeutig bestimmen mit

$$b(z) = a(z) \cdot q(z) + r(z) \quad \text{und} \quad \deg r(z) < \deg a(z).$$

Damit hat man

$$\frac{b(z)}{a(z)} = q(z) + \frac{r(z)}{a(z)},$$

d.h. bis auf das zusätzliche Polynom $q(z)$ hat man genau die Verhältnisse wie im obigen Theorem.

3.10 Newtons Formel und die Folgen

$\mathbb{C}[X]_{\leq k}$ bezeichnet der Vektorraum der (komplexen) Polynome vom Grad $\leq k$. Dieser Vektorraum hat viele verschiedene Basen, z.B.

- Standardbasis $1, X, X^2, \dots, X^k$
- Newton-Basis

$$(X)_j = X(X-1) \cdots (X-j+1) \quad (0 \leq j \leq k)$$

- Binomialbasis

$$\binom{X}{j} = \frac{(X)_j}{j!} \quad (0 \leq j \leq k)$$

- Euler-Basis

$$[X]_{k,j} = X^j(1-X)^{k-j} \quad (0 \leq j \leq k)$$

Dass diese auch wirklich Basen sind, sieht man leicht:

- Bei der Standardbasis ist das bereits mit der Definition von $\mathbb{C}[X]_{\leq k}$ festgelegt.

- Zwischen der Standardbasis und der Newton-Basis besteht folgende Beziehung (Basistransformation):

$$\begin{bmatrix} (X)_0 \\ (X)_1 \\ (X)_2 \\ (X)_3 \\ (X)_4 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 & & & & & \\ 0 & 1 & & & & \\ 0 & -1 & 1 & & & \\ 0 & 2 & -3 & 1 & & \\ 0 & -6 & 11 & -6 & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} 1 \\ X \\ X^2 \\ X^3 \\ X^4 \\ \vdots \end{bmatrix}$$

Die Zahlenwerte in der Matrix sind als *Stirling-Zahlen erster Art* in der Literatur bekannt und bestens untersucht. Die Matrix ist offensichtlich invertierbar:

$$\begin{bmatrix} 1 \\ X \\ X^2 \\ X^3 \\ X^4 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 & & & & & \\ 0 & 1 & & & & \\ 0 & 1 & 1 & & & \\ 0 & 1 & 3 & 1 & & \\ 0 & 1 & 7 & 6 & 1 & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} (X)_0 \\ (X)_1 \\ (X)_2 \\ (X)_3 \\ (X)_4 \\ \vdots \end{bmatrix}$$

Die Zahlenwerte in dieser Matrix sind als *Stirling-Zahlen zweiter Art* in der Literatur bekannt und bestens untersucht.

- Die Binomialbasis geht aus der Newton-Basis durch Multiplikation mit Skalaren hervor.
- Zwischen der Standardbasis und der Euler-Basis besteht folgende Beziehung (Basistransformation):

$$\begin{bmatrix} [X]_{k,k} \\ [X]_{k,k-1} \\ [X]_{k,k-2} \\ [X]_{k,k-3} \\ \vdots \\ [X]_{k,0} \end{bmatrix} = \begin{bmatrix} 1 & & & & & \\ -1 & 1 & & & & \\ 1 & -2 & 1 & & & \\ -1 & 3 & -3 & 1 & & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \\ \pm \binom{k}{k} & \pm \binom{k}{k-1} & \pm \binom{k}{k-2} & \pm \binom{k}{k-3} & \cdots & \binom{k}{0} \end{bmatrix} \begin{bmatrix} X^k \\ X^{k-1} \\ X^{k-2} \\ X^{k-3} \\ \vdots \\ 1 \end{bmatrix}$$

und auch diese Beziehung lässt sich umkehren:

$$\begin{bmatrix} X^k \\ X^{k-1} \\ X^{k-2} \\ X^{k-3} \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & & & & & \\ 1 & 1 & & & & \\ 1 & 2 & 1 & & & \\ 1 & 3 & 3 & 1 & & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \\ \binom{k}{k} & \binom{k}{k-1} & \binom{k}{k-2} & \binom{k}{k-3} & \cdots & \binom{k}{0} \end{bmatrix} \begin{bmatrix} [X]_{k,k} \\ [X]_{k,k-1} \\ [X]_{k,k-2} \\ [X]_{k,k-3} \\ \vdots \\ [X]_{k,0} \end{bmatrix}$$

Als Folgerung aus diesen Aussagen erhält man:

Satz 14. *Die Abbildung*

$$p(X) = \alpha_0 \binom{X}{0} + \cdots + \alpha_k \binom{X}{k} \mapsto \widehat{p}(X) = \alpha_0 [X]_{k,0} + \cdots + \alpha_k [X]_{k,k}$$

ist eine bijektive lineare Transformation, also eine Basistransformation des Vektorraums $\mathbb{C}[X]_k$.

Betrachte nun die im Einheitskreis $|z| < 1$ konvergierende) unendliche geometrische Reihe

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

Durch k -fache Ableitung erhält man aus

$$\frac{1}{k!} \left(\frac{d}{dz} \right)^k \frac{1}{1-z} = \frac{1}{k!} \left(\frac{d}{dz} \right)^k \sum_{n \geq 0} z^n$$

die Beziehung

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

die man besser noch in der Form

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

schreibt. Die ist (in heutiger Schreibweise) *Newtons Binomialreihe*. Sie spielt eine Rolle in der folgenden Umformung, die auf den vorigen Satz Bezug nimmt.

Ist $p(X) \in \mathbb{C}[X]_{\leq k}$ ein Polynom, so gilt

$$\begin{aligned} \sum_{n \geq 0} p(n) z^n &= \sum_{n \geq 0} \left(\sum_{j=0}^k \alpha_j \binom{n}{j} \right) z^n \\ &= \sum_{j=0}^k \alpha_j \left(\sum_{n \geq 0} \binom{n}{j} z^n \right) \\ &= \sum_{j=0}^k \alpha_j \frac{z^j}{(1-z)^{j+1}} \\ &= \frac{\sum_{j=0}^k \alpha_j z^j (1-z)^{k-j}}{(1-z)^{k+1}} = \frac{\widehat{p}(z)}{(1-z)^{k+1}} \end{aligned}$$

Satz 15. Für $p(X) \in \mathbb{C}[X]_{\leq k}$ gilt die Reihenentwicklung

$$\sum_{n \geq 0} p(n) z^n = \frac{\widehat{p}(z)}{(1-z)^{k+1}}$$

die für $|z| < 1$ konvergiert.

Seien nun $a(X), b(X)$ Polynome, wobei $a(0) = 1$ sein soll und beide keinen gemeinsamen Teiler (=keine gemeinsame Nullstelle) haben sollen. Sei

$$a(z) = \prod_{j=1}^{\ell} (1 - \lambda_j z)^{t_j}$$

die Zerlegung $a(z)$ in Linearfaktoren, d.h. die λ_j sind die *verschiedenen* Nullstellen von $a(X)$ und die t_j ihre Vielfachheiten. Es sei

$$f(z) = \sum_{n \geq 0} f_n z^n = \frac{b(z)}{a(z)}$$

die durch $\langle a(X), b(X) \rangle$ bestimmte strikte rationale Funktion.

Satz 16. Es gibt eindeutig bestimmte Polynome $p_1(X), \dots, p_\ell(X)$, wobei p_j einen Grad $< t_j$ hat, mit

$$f_n = \sum_{j=1}^{\ell} p_j(n) \lambda_j^n \quad (n \geq 0)$$

Das ergibt sich nun aus dem vorigen Satz und der Partialbruchzerlegung

$$\begin{aligned} \sum_{n \geq 0} f_n z^n &= \frac{b(z)}{a(z)} = \frac{b(z)}{\prod_{j=1}^{\ell} (1 - \lambda_j z)^{t_j}} \\ &= \sum_{j=1}^{\ell} \frac{\widehat{p}_j(z)}{(1 - \lambda_j z)^{t_j}} = \sum_{j=1}^{\ell} \sum_{n \geq 0} p_j(n) \lambda_j^n z^n \\ &= \sum_{n \geq 0} \sum_{j=1}^{\ell} p_j(n) \lambda_j^n z^n \end{aligned}$$

durch Koeffizientenvergleich.

3.11 Inhomogene (forcierte) lineare Rekursionen

Der Formalismus der rationalen Funktionen erlaubt es in wichtigen Fällen, die Untersuchung *inhomogener* (forcierter) Rekursionen auf die Untersuchung *homogener* Rekursionen zurückzuführen.

Betrachten wir die homogene C-Rekursion

$$(*) \quad x_n = a_1 x_{n-1} + a_2 x_{n-2} + \cdots + a_k x_{n-k} \quad (n \geq k)$$

definiert durch das Rekursionspolynom

$$a(z) = 1 - a_1 z - a_2 z^2 - \cdots - a_k z^k$$

und die inhomogene C-Rekursion

$$(**) \quad x'_n = a_1 x'_{n-1} + a_2 x'_{n-2} + \cdots + a_k x'_{n-k} + y_{n-k} \quad (n \geq k),$$

bei der die Folge $\mathbf{y} = (y_n)_{n \geq 0}$ eine "von aussen" einwirkende Grösse modelliert, wie durch die grafische Darstellung als Schieberegister deutlich wird:

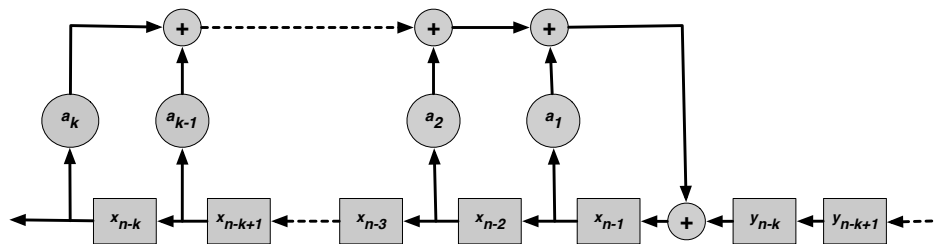


Abbildung 4: Forciertes Schieberegister

Die Folge \mathbf{y} wird auch durch ihre Potenzreihe

$$y(z) = y_0 + y_1 z + y_2 z^2 + \cdots = \sum_{n \geq 0} y_n z^n$$

vertreten.

Ein Lösung $\mathbf{x} = (x_n)_{n \geq 0}$ der Rekursion (*) ist gegeben durch ein Polynom $b(z)$ mit einem Grad $< k$:

$$\sum_{n \geq 0} x_n z^n = \frac{b(z)}{a(z)}$$

Eine analoge Überlegung, die zu dieser Beziehung geführt hat, zeigt, dass eine Lösung $\mathbf{x}' = (x'_n)_{n \geq 0}$ durch

$$\sum_{n \geq 0} x'_n z^n = \frac{b(z) + z^k y(z)}{a(z)}$$

gegeben ist. Interessant ist insbesondere der Fall, wenn die Folge \mathbf{y} selbst eine C-rekursive Folge ist, also gegeben durch ein Paar $\langle c(z), d(z) \rangle$ von Polynomen. Dann gilt nämlich

$$\sum_{n \geq 0} x'_n z^n = \frac{b(z) c(z) + z^k d(z)}{a(z) c(z)}.$$

Daran sieht man, dass die Folge \mathbf{x}' einer *homogenen* Rekursion mit dem Rekursionspolynom $a(z) \cdot c(z)$ genügt! Man kann also zur Untersuchung der Folge \mathbf{x}' die homogene Theorie anwenden.

Bemerkung 6. Etwas Vorsicht ist geboten: die beiden Polynome $a'(z) = a(z) c(z)$ und $b'(z) = b(z) c(z) + z^k d(z)$ können durchaus gemeinsame Teiler haben! Das Paar $\langle a'(z), b'(z) \rangle$ ist also nicht notwendig schon die ‐ausgekürzte‐ Darstellung der rationalen Funktion $\sum_{n \geq 0} x'_n z^n$. Mit anderen Worten: die Rekursion für \mathbf{x}' , die durch das Polynom $a(z) c(z)$ gegeben ist, ist nicht notwendig die Rekursion kleinstmöglichen Grades.

Beispiel 3. Ein häufiger Fall ist der, dass die durch das Polynom $a(z)$ beschriebene Rekursion eine dominierende Nullstelle λ_1 hat. Ist dann speziell $\mathbf{y} = (\gamma^n)_{n \geq 0}$, so gilt $d(z) = 1$ und $c(z) = 1 - \gamma z$, und somit

$$\sum_{n \geq 0} x'_n z^n = \frac{b(z) (1 - \gamma z) + z^k}{a(z) (1 - \gamma z)}.$$

Jetzt sind drei Fälle möglich

$$\begin{aligned} \lambda_1 > \gamma &\Rightarrow x'_n \in \Theta(\lambda_1^n) \\ \lambda_1 = \gamma &\Rightarrow x'_n \in \Theta(n \cdot \lambda_1^n) \\ \lambda_1 < \gamma &\Rightarrow x'_n \in \Theta(\gamma^n) \end{aligned}$$

Situationen wie diese treten insbesondere im Kontext der Divide-and-Conquer-Rekursionen auf – siehe Kapitel 5.

3.12 Kommentare, Literatur, Ausblicke

3.12.1 C-rekursive Folgen, Differenzenrechnung

Systematisch gesehen gehören die C-rekursiven Folgen und die sie definierenden *Linearen Rekursionen mit konstanten Koeffizienten* in den Bereich der *Differenzenrechnung*, dem diskreten Analogon der Differentialrechnung. Ihre mathematische Theorie ist in vielen Texten ausführlich abgehandelt. Das Monografien [4] von EVEREST ET AL. und [1] (aus etwas anderer Perspektive) von ALLOUCHE und SHALLIT markieren anspruchsvolle und tiefeschürfende Abhandlung zum Thema der rekursiv definierten Folgen allgemein. Wesentlich elementarer, dafür mit vielen instruktiven Beispielen versehen, ist der Text [3] von ELAYDI. Eine schöne elementare Darstellung der Techniken rund um die linearen Rekursionsgleichungen bietet Kapitel 10 in dem Buch [9].

3.12.2 Schieberegisterfolgen, Pseudo-Zufallsfolgen

Schiebergisterfolgen sind nichts anderes als eine technische Realisierung von C-Rekursionen, wobei meist nicht über dem K -Körper der komplexen Zahlen gerechnet wird, sondern über eine *endlichen* Körper, vorzugsweise dem zweielementigen Körper \mathbb{F}_2 . Die systematische Untersuchung C-rekursiver Folgen unter diesem Aspekt wurde in [11, 12] von ZIERLER begründet. Eine erste Zusammenfassung ist der "Klassiker" [5] von GOLOMB. Bücher über *Algebraische Codierungstheorie* wie z.B. [2] von BERLEKAMP, ein anderer "Klassiker", machen davon ausgiebig Gebrauch. Eine umfassende Darstellung findet man auch im Kapitel 8 des Monumentalwerks [7] von LIDL und NIEDERREITER.

3.12.3 Algorithmus von Berlekamp-Massey

Das Problem, zu einer gegebenen C-rekursiven Folge die Rekursionskoeffizienten zu bestimmen, hier als *reverse engineering* bezeichnet, ist eine fundamentale Aufgabe in der Nachrichtentechnik. Besonders hervorzuheben ist das Gebiet der *Fehlerkorrigierenden Codes (error-correcting codes)*, bei der die effizienten Decodierungsalgorithmen für die sog. verallgemeinerten REED-SOLOMON-Codes, der wichtigsten Klasse von fehlerkorrigierenden Codes überhaupt auf der Lösung genau dieser Rekonstruktionsaufgabe beruhen. Der prominenteste Algorithmus dafür stammt von E. BERLEKAMP (siehe [2]) und J. MASSEY (siehe [8]). In dem sehr gründlichen Monografie [10] von ROTH findet man eine aktuelle Darstellung all dieser Zusammenhänge.

3.12.4 z -Transformation, Lineare Differenzen- und Differentialgl.

In der linearen Systemtheorie/Nachrichtentechnik spielen Schieberegister und die von ihnen erzeugten Folgen eine besondere Rolle. Allerdings ist die Terminologie eine andere. Systemtheoretiker sprechen von der z -Transformation einer Folge oder eines "Signals", was bis auf fachbedingte Traditionsunterschiede in der Notation nichts anders ist als Beziehung zwischen einer Koeffizientenfolge $(f_n)_{n \in \mathbb{N}}$ und der Potenzreihe $f(z) = \sum_{n \geq 0} f_n z^n$. Wie dargestellt, gilt dabei i.w. eine 1-1-Beziehung

C-rekursive Folgen \leftrightarrow Rationale Funktionen.

Wer sich schon einmal mit *linearen Differentialgleichungen mit konstanten Koeffizienten* befasst hat, wird feststellen, dass die Theorie der Lösungen solcher Gleichungen ganz analog zur hier behandelten Theorie der Lösungen von C-Rekursionen aufgebaut ist. Auch dabei spielen rationale Funktionen eine zentrale Rolle und als "workhorse" tritt in diesem Fall an Stelle der z -Transformation die LAPLACE-Transformation.

Lösungen linearer Dgln. \leftrightarrow Rationale Funktionen.

Hinter beiden Komplexen, *C-Rekursionen* und *Lineare Differentialgleichungen* steckt also die gleiche lineare Theorie!

Auch hierfür gibt es eine umfangreiche Ingenieursliteratur, der Titel [6] von GRAF ist nur ein Beispiel.

Literatur

- [1] Jean-Paul Allouche and Jeffrey Shallit. *Automatic Sequences*. Cambridge University Press, Cambridge, 2003. Theory, applications, generalizations.
- [2] E.R. Berlekamp. *Algebraic Coding Theory*. Aegean Park Press, 1984.
- [3] Saber N. Elaydi. *An Introduction to Difference Equations*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1999.
- [4] Graham Everest, Alf van der Poorten, Igor Shparlinski, and Thomas Ward. *Recurrence Sequences*, volume 104 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2003.
- [5] Solomon W. Golomb. *Shift Register Sequences*. Aegean Park Press, 1982.

-
- [6] Urs Graf. *Applied Laplace Transforms and z-Transforms for Scientists and Engineers*. Birkhäuser Verlag, Basel, 2004. A computational approach using a *Mathematica* package, With 1 CD-ROM (Windows and LINUX).
- [7] Rudolf Lidl and Harald Niederreiter. *Finite Fields*, volume 20 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Company Advanced Book Program, Reading, MA, 1983. With a foreword by P. M. Cohn.
- [8] James L. Massey. Shift-register synthesis and BCH decoding. *IEEE Trans. Information Theory*, IT-15:122–127, 1969.
- [9] Robert J. McEliece, Robert B. Ash, and Carol Ash. *Introduction to Discrete Mathematics*. McGraw-Hill, 1989.
- [10] Ron M. Roth. *Introduction to Coding Theory*. Cambridge University Press, 2006.
- [11] Neal Zierler. Linear recurring sequences. *J. Soc. Indust. Appl. Math.*, 7:31–48, 1959.
- [12] Neal Zierler. Linear recurring sequences and error-correcting codes. In *Error Correcting Codes (Proc. Sympos. Math. Res. Center, Madison, Wis., 1968)*, pages 47–59. Wiley, New York, 1968.