Kopf-Adler-Muster in Münzwurfserien, unendliche Reihen und Fibonacci-Zahlen

HANS HUMENBERGER, DORTMUND

Zusammenfassung: *Wir wollen hier Serien von Münzwürfen näher beleuchten, also z. B. Serien der Gestalt $KAKKAKAAAK$...... (K steht für „Kopf“ und A für „Adler“). Insbesondere soll es um die zweigliedrigen Muster bzw. Wörter KA bzw. KK gehen. Es wird sich herausstellen, daß man im Durchschnitt 6-mal die Münze werfen muß, um das Muster KK (zweimal hintereinander K) zu erhalten, wohingegen für KA im Durchschnitt 4 Versuche reichen (die Existenz aller in Rede stehenden Erwartungswerte wird nicht bewiesen, sondern als intuitiv einsichtig vorausgesetzt). Es wird sich aber auch herausstellen, daß die Wahrscheinlichkeit, daß in einer Wurfserie KK früher als KA erscheint (oder umgekehrt), trotzdem $\frac{1}{2}$ beträgt. Schließlich wird gezeigt, wie bei der Anzahl KK-freier Serien einer gewissen Länge n Fibonacci-Zahlen eine Rolle spielen.*

Es gibt mehrere Möglichkeiten, sich Gedanken zum Thema „Muster in Münzwurfserien“ zu machen, so z. B. mit Hilfe bedingter Erwartungswerte – siehe z. B. (Humenberger 2000). Wir wollen hier jedoch bewußt auf diese verzichten und die Methode „direkte Berechnung des Erwartungswertes“ (gemeint ist mittels dessen Definition) anwenden, obwohl diese manchem vielleicht umständlich erscheinen mag.

Weitere Methoden zur Behandlung von Mustern in binären Sequenzen wären z. B. gegeben durch: (1) Graphen und „Mittelwertsregeln“ – siehe (Engel 1976, S.18ff insbesondere S. 22–26); (2) mit Hilfe des sogenannten Conway-Algorithmus – beschrieben z. B. in (Gardner 1974, S.123) oder (Szekely 1990, S.62); dieser ist zwar sehr „verblüffend“ und auch sehr leicht durchzuführen, aber relativ schwierig einzusehen bzw. zu begründen – z. B. (Li 1980).

Muster in Serien und ihre Wahrscheinlichkeiten finden in der Genetik („*Sequence-Matching*“) eine wichtige außermathematische Anwendung – siehe z. B. (Waterman 1995).

1 Unendliche Reihen

Der „mathematische Hintergrund“ folgender Überlegungen sind unendliche (geometrische) Reihen. Er kann durch die bekannte Summenformel

$$\boxed{\sum_{n=1}^{\infty} x^n = \frac{x}{1-x}} \qquad \text{für} \quad |x| < 1 \tag{1}$$

und z. B. durch folgende Identität („gliedweises Differenzieren“):

$$\boxed{\sum_{n=1}^{\infty} n\, x^{n-1}} = \left(\sum_{n=1}^{\infty} x^n\right)' = \left(\frac{x}{1-x}\right)' \tag{2}$$

$$= \boxed{\frac{1}{(1-x)^2}} \tag{3}$$

(ebenfalls für $|x| < 1$) beschrieben werden. Für $x = \frac{1}{2}$ ergibt sich z. B.

$$\boxed{\sum_{n=1}^{\infty} n\, \frac{1}{2^{n-1}}} = 1 + 2 \cdot \frac{1}{2^1} + 3 \cdot \frac{1}{2^2} + \ldots \tag{4}$$

$$= \frac{1}{\left(1-\frac{1}{2}\right)^2} = \boxed{4}\,. \tag{5}$$

Für die Begründung der Identität (3) wäre Differentialrechnung (gliedweises Differenzieren) auch vermeidbar (sollte z. B. Differentialrechnung noch nicht zur Verfügung stehen oder aus anderen Gründen vermieden werden). Aus

$$S := \sum_{n=1}^{\infty} n\, x^{n-1} = 1 + 2x + 3x^2 + 4x^3 + 5x^4 + \ldots$$

erhält man für $|x| < 1$ durch Multiplikation mit $(1-x)$

$$\begin{aligned}(1-x)S &= (1+2x+3x^2+4x^3+\ldots) - \\ &\quad -(x+2x^2+3x^3+4x^4+\ldots) \\ &= 1+x+x^2+x^3+\cdots = \frac{1}{1-x}\end{aligned}$$

[das „Umordnen“ ist wegen der absoluten Konvergenz erlaubt; man könnte auch etwas formaler zunächst nur endliche Summen und dann den Grenzwert bilden] und daraus sofort

$$\sum_{n=1}^{\infty} n\,x^{n-1} = S = \frac{1}{(1-x)^2} \quad (|x|<1)\,. \tag{6}$$

Definition: Es sei X eine diskrete Zufallsvariable mit $\{x_1, x_2, x_3, \ldots, x_k, \ldots\}$ als mögliche Werte. Unter dem Erwartungswert $E(X)$ der Zufallsvariable X versteht man die Summe

$$\boxed{E(X) \stackrel{\text{def}}{=} \sum_k x_k\, P(X = x_k)\,,}$$

falls diese (möglicherweise unendliche) Reihe absolut konvergiert.

Mit Hilfe von (3) bzw. (6) kann z. B. der Erwartungswert einer *geometrisch verteilten Zufallsvariable* („Wartezeit, bis bei einem Bernoulli-Versuch ein Ereignis A, das bei jeder Versuchswiederholung mit Wahrscheinlichkeit p auftritt, *erstmalig* eintritt“) berechnet werden. Dies ist beim ersten Versuch eben mit Wahrscheinlichkeit p der Fall, beim zweiten mit Wahrscheinlichkeit $(1-p)\,p$, beim dritten mit Wahrscheinlichkeit $(1-p)^2\,p$ ..., beim n-ten Versuch mit Wahrscheinlichkeit $(1-p)^{n-1}\,p$ („bei den ersten $n-1$ Versuchen *nicht A*, beim n-ten aber A“). Für den zugehörigen Erwartungswert der Wartezeit (= Anzahl der nötigen Versuche) bis A erstmalig auftritt, erhalten wir daher

$$\begin{aligned}E(X) &= 1\cdot p + 2\cdot(1-p)\,p + 3\cdot(1-p)^2\,p + \\ &\quad + \cdots + n\cdot(1-p)^{n-1}\,p + \ldots\end{aligned}$$

bzw.

$$\begin{aligned}\boxed{E(X)} &= p\sum_{n=1}^{\infty} n\,(1-p)^{n-1} && (7)\\ &= p\cdot\frac{1}{(1-(1-p))^2} = \boxed{\frac{1}{p}}\,. && (8)\end{aligned}$$

Man braucht also im Durchschnitt zwei Versuche ($p = \frac{1}{2}$), damit beim wiederholten Münzwurf K erscheint, bzw. 6 Versuche ($p = \frac{1}{6}$), damit beim wiederholten Würfeln eine *Sechs* erscheint.

Der Erwartungswert (8) kann auch ohne Rückgriff auf (3) bzw. (6) hergeleitet werden (analog zur obigen Methode ohne Differentialrechnung). Wir bezeichnen der Übersichtlichkeit halber $1 - p \stackrel{\text{def}}{=} q$ und erhalten

$$\begin{aligned}&E(X) = \\ &= 1\cdot p + 2\cdot q\,p + \cdots + n\cdot q^{n-1}\,p + \ldots \\ &= (1-q)\cdot\left(1+2q + \cdots + nq^{n-1} + \ldots\right) \\ &= \left(1+2q + 3q^2 + \cdots + nq^{n-1} + \ldots\right) \\ &- \left(q + 2q^2 + 3q^3 + \ldots + (n-1)q^{n-1} + \ldots\right) \\ &= 1 + q + q^2 + q^3 + q^4 + \cdots = \frac{1}{1-q} = \frac{1}{p}\end{aligned}$$

Durch *zwei*maliges gliedweises Differenzieren der geometrischen Reihe erhält man eine weitere wichtige Identität, auf die wir später noch einmal zurückgreifen werden (mit ihrer Hilfe kann z. B. die Varianz einer geometrisch verteilten Zufallsvariable berechnet werden, vgl. z. B. (Reichel/-Hanisch/Müller 1992, S.166f):

$$\begin{aligned}\boxed{\sum_{n=2}^{\infty} n\,(n-1)\,x^{n-2}} &= \left(\sum_{n=1}^{\infty} x^n\right)'' && (9)\\ &= \left(\frac{x}{1-x}\right)'' && (10)\\ &= \boxed{\frac{2}{(1-x)^3}} && (11)\end{aligned}$$

Für $x = \frac{1}{2}$ erhalten wir damit

$$\begin{aligned}\boxed{\sum_{n=2}^{\infty} n\,\frac{n-1}{2^n}} &= \frac{1}{4}\cdot\sum_{n=2}^{\infty} n\,\frac{n-1}{2^{n-2}} && (12)\\ &= \frac{1}{4}\cdot\frac{2}{\left(1-\frac{1}{2}\right)^3} = \boxed{4} && (13)\end{aligned}$$

Für die Beziehung (11) weisen wir nur darauf hin, daß durch Multiplikation der interessierenden Summe mit $(1-x)$ (analog zu oben) und Heranziehung von (6) Differentialrechnung zur Bestimmung des Summenwertes vermeidbar wäre.

2 Gleichwahrscheinlichkeit von „KA früher als KK“ und „KK früher als KA“ in dieser Sichtweise

Mit „KA früher als KK“ ist gemeint: KA erscheint früher als KK als Ergebnis von zwei unmittelbar aufeinanderfolgenden Würfen.

Daß die beiden Ereignisse „KA früher als KK“ und „KK früher als KA“ gleichwahrscheinlich sind (nämlich jeweils $\frac{1}{2}$) kann ganz einfach begründet werden: nach dem ersten K kommt mit gleicher Wahrscheinlichkeit entweder ein weiteres K oder ein A (eben mit Wahrscheinlichkeit $\frac{1}{2}$) und damit tritt genau eines der beiden genannten Ereignisse ein.

Bemerkung: Ein erstes K kommt mit Wahrscheinlichkeit 1 *irgendwann einmal* vor (d. h. in einer „unendlich langen“ Serie von Münzwürfen; man kann sogar relativ leicht zeigen, daß *jedes endliche Muster* mit Wahrscheinlichkeit 1 in einer unendlich langen Serie vorkommt). *Keines* der beiden genannten Muster würde nur bei $AAA\ldots AA\ldots$ (immer nur As) eintreten, ein Ereignis mit Wahrscheinlichkeit 0.

Die Wahrscheinlichkeiten $\frac{1}{2}$ für „KA früher als KK“ und „KK früher als KA“ können aber auch anders einsichtig gemacht werden, indem wir überlegen, wie viele Versuchsausgänge das jeweilige Ereignis herbeiführen („günstige Möglichkeiten“) – wir wollen hier absichtlich mit unendlichen Reihen arbeiten und möglicherweise einfachere Lösungen eher zurückdrängen – dafür sei z. B. auf (Humenberger 2000) verwiesen. Die folgende (kompliziertere) Sichtweise bringt jedoch gegenüber der schon erwähnten Begründung für die Gleichwahrscheinlichkeit von „KK früher als KA“ und „KA früher als KK“ (nach dem ersten K kommt mit gleicher Wahrscheinlichkeit ein K bzw. ein A) die Erkenntnis, daß es bei jedem Muster der Länge n genau eine für das jeweilige Ereignis „günstige“ Möglichkeit gibt.

1. Das Ereignis $\boxed{\text{„KK früher als KA“}}$ tritt genau dann bei einem bestimmten n ein, wenn beim $n-1$-ten und n-ten Wurf zusammen *erstmalig* KK steht und bis dahin kein KA vorgekommen ist. Jedes $n \geq 2$ kommt dafür prinzipiell in Frage. Wie viele Möglichkeiten gibt es für *solche* („günstige“) Serien der Länge n?

   Bei $n = 2$ gibt es dafür nur eine Möglichkeit, nämlich KK; bei $n = 3$ ebenfalls nur eine $(A|KK)$ und auch bei jedem anderen n nur *genau eine Möglichkeit* $(A\ldots AA|KK)$, da vor dem letzten KK weder ein anderes KK noch ein KA auftreten darf. An der $(n-2)$-ten Stelle muß also ein A stehen, denn sonst wäre vor dem in Rede stehenden KK an der Stelle n bereits ein KK bei $n-1$ eingetreten, und vor diesem A an der drittletzten Stelle dürfen natürlich nur As stehen (sonst wäre KA bereits *vorher* aufgetreten).

   Da für jedes $n \geq 2$ die Zahl der möglichen „KA-Serien der Länge n“ 2^n und die Zahl der „günstigen“ 1 beträgt, erhalten wir für die Gesamtwahrscheinlichkeit des Ereignisses „KK früher als KA“ (unendliche geometrische Reihe, siehe (1))

$$\boxed{P(KK \text{ früher als } KA)} =$$

$$= \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \cdots = \sum_{n=2}^{\infty} \frac{1}{2^n} = \boxed{\frac{1}{2}}.$$

2. $\boxed{\text{„}KA \text{ früher als } KK\text{“}}$ Analog gibt es auch für das Ereignis „erstmalig KA an der Stelle n, ohne daß vorher KK aufgetreten ist“ bei jedem $n \geq 2$ jeweils genau eine Möglichkeit, nämlich wieder nur A's vorher $(A\ldots AA|KA)$, denn vor dem K an der vorletzten Stelle muß wieder A stehen – wie oben. Wir erhalten für die Gesamtwahrscheinlichkeit des Ereignisses „KA früher als KK“ ebenfalls

$$\boxed{P(KA \text{ früher als } KK)} =$$

$$= \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \cdots = \sum_{n=2}^{\infty} \frac{1}{2^n} = \boxed{\frac{1}{2}}.$$

3 Die Erwartungswerte für ein beliebiges Doppelmuster: „KK oder AA“, und für eine beliebige Abwechslung: „AK oder KA“

$\boxed{E(KK \text{ oder } AA)}$ – die durchschnittliche Wartezeit auf das erste (beliebige) „Doppelmuster“.

Das *erste* Doppelmuster in einer Serie kann prinzipiell bei jedem $n \geq 2$ realisiert werden – dies sind die möglichen Ausprägungen der zugehörigen Zufallsvariable, deren Erwartungswert wir berechnen:

$$E(KK \text{ oder } AA) = \sum_{n=2}^{\infty} n \cdot p_n,$$

wobei p_n die Wahrscheinlichkeit ist, daß das erste Doppelmuster nach n Würfen auftritt; wie groß sind die zugehörigen Wahrscheinlichkeiten p_n? Wir werden sie wieder mittels „günstig/möglich“ (mit „möglich“ $= 2^n$) berechnen.

Man überlegt sich leicht, daß es für jedes $n \geq 2$ *genau zwei* Möglichkeiten gibt, daß gerade an dieser Stelle erstmalig ein Doppelmuster (KK oder AA) auftritt: Da es vorher immer *Abwechslungen* gegeben haben muß, kann das Gesamtmuster bei KK als erstmaliges Doppelmuster an der n-ten Stelle nur so aussehen: $\ldots KAKA|KK$ und bei AA als erstmaliges Doppelmuster an der n-ten Stelle nur so: $\ldots AKAK|AA$. Bei n Würfen gibt es jeweils 2^n mögliche n-Serien, so daß wir mit (5)

$$\begin{aligned} \boxed{E(KK \text{ oder } AA)} &= \sum_{n=2}^{\infty} n \frac{2}{2^n} \\ &= \sum_{n=2}^{\infty} \frac{n}{2^{n-1}} \\ &= 4 - 1 = \boxed{3} \end{aligned}$$

erhalten. (Die Subtraktion $\cdots - 1$ am Schluß, weil die Summation erst bei $n = 2$ beginnt.)

$\boxed{E(AK \text{ oder } KA)}$ – die durchschnittliche Wartezeit auf die erste (beliebige) „Abwechslung“.

Analog muß bei KA als erstmalige Abwechslung das Muster die Gestalt $KK \ldots KK|KA$ haben, bzw. $AA \ldots AA|AK$ bei AK als erstmalige Abwechslung, also auch *genau zwei* Möglichkeiten für jedes $n \geq 2$, daß gerade an dieser Stelle erstmalig eine Abwechslung stattfindet. Es ergibt sich daher wieder

$$\begin{aligned} \boxed{E(AK \text{ oder } KA)} &= \sum_{n=2}^{\infty} n \frac{2}{2^n} \\ &= \sum_{n=2}^{\infty} \frac{n}{2^{n-1}} \\ &= 4 - 1 = \boxed{3}. \end{aligned}$$

Als nächstes wollen wir uns dem Problem widmen, wie oft man im Mittel eine Münze werfen muß, damit das Muster KA erstmalig erscheint, also die Berechnung von $E(KA)$, hier mit Hilfe der Identität (11).

4 E(KA) = 4 durch Überlegungen zu KA-freien Serien

Die Erwartungswerte der nötigen Würfe für eine *beliebige Abwechslung* (also „AK oder KA“) bzw. für ein *beliebiges Doppelmuster* (also „KK oder AA“) sind jeweils 3. Dadurch könnte man sich zur Vermutung hinreißen lassen, daß auch die zugehörigen Erwartungswerte für ein *bestimmtes Doppelmuster* (also z. B. KK) bzw. für ein *bestimmtes* „Abwechslungsmuster“ (z. B. KA) gleich sind. Wir werden jedoch sehen, daß dies keineswegs der Fall ist!

Sei X die Zufallsvariable, die zählt, wann KA zum ersten Mal auftritt. Für die Berechnung von

$$E(KA) = E(X) = \sum_{n=2}^{\infty} n\, P(X = n)$$

stellen wir zunächst Überlegungen für die in dieser Summe auftretenden Wahrscheinlichkeiten $P(X = n)$ bzw. für die Anzahl der jeweils „günstigen Fälle“ x_n an; die Anzahl der möglichen Wurfserien der Länge n beträgt bei jedem n selbstverständlich wieder 2^n, so daß $P(X = n) = \frac{x_n}{2^n}$ ist.

Für $n = 2$ gibt es trivialerweise eine einzige Möglichkeit, nämlich KA selbst ($x_2 = 1$). Für $n = 3$ gibt es zwei Möglichkeiten, daß KA am Schluß erstmalig auftritt: $K|KA$ und $A|KA$, d.h.

$x_3 = 2$. Bei einem viergliedrigen Muster gibt es drei für die beschriebene Situation günstige Fälle: $AA|KA$, $AK|KA$, und $KK|KA$ ($x_4 = 3$, an den ersten beiden Stellen darf KA nicht vorkommen). Bei $n = 5$ müssen wir analog überlegen, wie viele Möglichkeiten es gibt, die ersten drei Stellen „KA-frei" zu besetzen; es gibt $x_5 = 4$ Möglichkeiten: $AAA|KA$, $AAK|KA$, $AKK|KA$ und $KKK|KA$.

Es bleibt also die Vermutung $\boxed{x_n = n - 1}$ zu begründen, daß es bei einer n-gliedrigen Serie $n - 1$ Möglichkeiten gibt, die ersten $n - 2$ Plätze „KA-frei" zu besetzen (KA soll ja an den letzten beiden Stellen zum ersten Mal auftauchen!). Dies ist relativ leicht zu begründen: Ein Kandidat einer KA-freien Serie ist jene, die ausschließlich aus A's besteht. Wenn ein A in dieser Serie durch ein K ersetzt wird, so müssen auch alle folgenden Stellen durch ein K ersetzt werden, denn sonst würde ja „vorzeitig" ein KA auftreten. Die einzigen Konkurrenten der reinen A-Serie sind demnach jene Serien, die ab irgendeiner Stelle nur mehr noch K's aufweisen, und solche gibt es bei einer Länge von $n - 2$ – die letzten beiden sollen ja KA sein – gerade $n - 2$ viele, insgesamt (mit der reinen A-Serie) also $n - 1$:

$$\underbrace{AAA\ldots AAA}_{n-2}|KA, \quad \underbrace{AAA\ldots AAK}_{n-2}|KA,$$

$$\underbrace{AAA\ldots AKK}_{n-2}|KA, \quad \underbrace{AAA\ldots KKK}_{n-2}|KA,$$

$$\ldots\ldots\ldots, \quad \underbrace{AAK\ldots KKK}_{n-2}|KA,$$

$$\underbrace{AKK\ldots KKK}_{n-2}|KA, \quad \underbrace{KKK\ldots KKK}_{n-2}|KA.$$

Wir erhalten daher für die jeweiligen Wahrscheinlichkeiten $P(X = n) = \frac{x_n}{2^n} = \frac{n-1}{2^n}$ und schließlich mit (13) für den zu berechnenden Erwartungswert

$$\boxed{E(KA)} = E(X) = \sum_{n=2}^{\infty} n\,\frac{n-1}{2^n} = \boxed{4}.$$

5 KK-freie Serien und Fibonacci-Zahlen

5.1 E(KK) = 6 – wie kommen Fibonacci-Zahlen hier herein?

Genau wie es zur Bestimmung von $E(KA)$ wichtig war, die Anzahl der KA-freien Serien der Länge n (bzw. $n-2$) zu kennen, so benötigen wir für die Berechnung von $E(KK)$ in dieser Sichtweise die Anzahl der KK-freien Serien gegebener Länge – wir werden sehen, daß diese i. a. viel höher ist!

Sei Y die Zufallsvariable, die zählt, wann KK erstmalig auftritt (dies ist analog ab $n = 2$ möglich). Zur Berechnung von

$$E(KK) = E(Y) = \sum_{n=2}^{\infty} n\,P(Y = n)$$

stellen wir erneut zunächst Überlegungen für die auftretenden Wahrscheinlichkeiten $P(Y = n)$ bzw. für die Anzahl der jeweils „günstigen Fälle" y_n an; die Anzahl der möglichen Fälle (Wurfserien der Länge n) beträgt bei jedem n wiederum 2^n, also $P(Y = n) = \frac{y_n}{2^n}$.

Für jedes $n \geq 2$ interessieren wir uns für die Anzahl der Möglichkeiten, die ersten $n - 2$ Plätze KK-frei zu belegen: $\underbrace{\ldots\ldots\ldots}_{n-2 \text{ Stellen}}|KK$. Bei $n = 2$ gibt es eine „günstige" Möglichkeit, KK selbst ($y_2 = 1$). Bei $n = 3$ gibt es ebenfalls nur eine Möglichkeit $A|KK$ ($y_3 = 1$). Im Fall $n = 4$ ergeben sich zwei günstige Möglichkeiten: $AA|KK$ und $KA|KK$ ($y_4 = 2$).

Bemerkung: Die $(n - 2)$-te Stelle – unmittelbar vor dem „|" – kann natürlich nur ein A sein, denn andernfalls gäbe es ein KK schon an der $(n - 1)$-ten Stelle und nicht – wie gefordert – erstmals an der n-ten Stelle; d. h. in jedem Fall n sind die letzten beiden Stellen durch KK und die drittletzte durch A „zu besetzen" und brauchen daher in die Überlegungen nicht (mehr) miteinbezogen zu werden. Die Frage nach den Möglichkeiten für „erstmalig KK an n-ter Stelle" ist also äquivalent mit der Frage nach der Anzahl der KK-freien Serien der Länge $n - 3$.

Für $n = 5$ erhalten wir $y_5 = 3$ Möglichkeiten, die ersten zwei Stellen KK-frei zu besetzen (die dritte muß A sein): $AA|A|KK$, $KA|A|KK$ und $AK|A|KK$. Für $n = 6$ ergeben sich $y_6 = 5$ „günstige" Möglichkeiten: $AAA|A|KK$, $KAA|A|KK$, $AKA|A|KK$, $AAK|A|KK$ und $KAK|A|KK$. Im Fall $n = 7$ – dies ergibt eine kurze Überlegung – sind es $y_7 = 8$ Möglichkeiten, die ersten vier Stellen KK-frei zu besetzen, z. B. $KAKA|A|KK$.

Die sich der Reihe nach (ab $n = 2$) ergebenden Anzahlen „günstiger" Möglichkeiten y_n (1, 1, 2, 3, 5, 8, ...) geben nun sicher Anlaß zur Vermutung, daß hinter dieser Zahlenfolge jene von Fibonacci steckt, was auch nicht schwierig zu zeigen ist.

Satz: f_n bezeichne die n-te Fibonacci-Zahl ($f_1 = 1$, $f_2 = 1$, $f_n \stackrel{\text{def}}{=} f_{n-1} + f_{n-2}$). Dann gibt es für jedes $n \geq 2$ genau $\boxed{f_{n-1}}$ Möglichkeiten, eine Serie der Länge n so zu besetzen, daß an der n-ten Stelle *erstmalig* KK auftritt, also Serien der Form $\underbrace{\ldots\ldots\ldots}_{n-3\text{ Stellen}}|A|KK$, wobei die ersten $n - 3$ Stellen KK-frei sein sollen.

Beweis: Sei $\boxed{y_n}$ die Anzahl der Möglichkeiten für eine Serie der gewünschten Form mit n Stellen: $\underbrace{\overset{K}{A}\ldots\ldots\ldots|A|KK}_{n\text{ Stellen}}$. Dann haben wir $\boxed{y_n = f_{n-1}}$ zu zeigen! Wir haben bereits überprüft, daß dies für die ersten Werte von $n \geq 2$ zutrifft: $y_2 = 1$, $y_3 = 1$, $y_4 = 2$, $y_5 = 3$, $y_6 = 5$. Zum Beweis der Allgemeingültigkeit unserer Vermutung (Behauptung) müssen wir „nur" noch zeigen, daß diese Werte y_n die Fibonacci-Rekursion

$$\boxed{y_n = y_{n-1} + y_{n-2}}$$

erfüllen.

Sei $\boxed{A_n \stackrel{\text{def}}{=}}$ ... Anzahl der Serien mit gewünschter Gestalt, die mit A beginnen: n Stellen, die ersten $n - 3$ Stellen KK-frei, also: $\underbrace{A\overset{K}{A}\ldots\ldots\ldots|A|KK}_{n-1\text{ Stellen}}$.

Analog $\boxed{K_n \stackrel{\text{def}}{=}}$ ... Anzahl der Serien mit gewünschter Gestalt, die mit K beginnen: n Stellen, die ersten $n - 3$ Stellen KK-frei, also: $KA!\underbrace{\overset{K}{A}\ldots\ldots\ldots|A|KK}_{n-2\text{ Stellen}}$. Dann ist klarerweise

$$\boxed{y_n = A_n + K_n\ .}$$

An der zweiten Stelle einer n-Serie, die mit A beginnt, kann – wie oben schon angedeutet – ein A oder ein K sein; daher muß $A_n = y_{n-1}$ gelten. An der zweiten Stelle einer n-Serie, die mit K beginnt, muß ein A sein (sonst wäre KK bereits realisiert); weil dadurch schon „zwei Stellen fixiert sind", muß $K_n = y_{n-2}$ gelten! Insgesamt erhalten wir dadurch die Behauptung für $n \geq 3$:

$$\boxed{y_n = A_n + K_n = y_{n-1} + y_{n-2}}$$

und daher

$$\boxed{y_n = f_{n-1}.}$$

Für den in Rede stehenden Erwartungswert $E(Y) = E(KK)$ bedeutet dieses Resultat zunächst $P(Y = n) = \frac{f_{n-1}}{2^n}$ und in weiterer Folge

$$\boxed{E(Y) = E(KK) = \sum_{n=2}^{\infty} n\,\frac{f_{n-1}}{2^n}}\ ,$$

wobei f_{n-1} für die $(n - 1)$-te Fibonacci-Zahl steht.

5.2 Die explizite Darstellung der Fibonacci-Zahlen

Die Tatsache $E(Y) = E(KK) = \sum_{n=2}^{\infty} n\,\frac{f_{n-1}}{2^n} = 6$ kann nun relativ leicht gezeigt werden, wenn z. B. eine explizite Darstellung von f_n zur Verfügung steht, die bekanntlich durch

$$\boxed{f_n = FIB(n) \stackrel{\text{def}}{=} \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]}$$

gegeben ist.

Wir führen – der Übersichtlichkeit halber – vorerst zwei naheliegende Abkürzungen ein, nämlich $\boxed{\alpha \stackrel{\text{def}}{=} \frac{1+\sqrt{5}}{2}}$ und $\boxed{\beta \stackrel{\text{def}}{=} \frac{1-\sqrt{5}}{2}}$.

Auch im Schulunterricht sollte *gezeigt* werden, daß obige Darstellung für f_n richtig ist! „Wie kann“ – könnte sonst ein allfälliger Einwand lauten – „so ein komplizierter Term für alle $n \in \mathbb{N}$ immer eine natürliche Zahl ergeben?“ – ein hier durchaus berechtigtes Unbehagen! Dafür geben wir zwei Beweismöglichkeiten an, wobei beim zweiten Beweis auch plausibel wird, *wie* man auf so einen kompliziert scheinenden Zusammenhang überhaupt kommen kann.

Beweis 1: Man überzeugt sich zunächst leicht (durch Einsetzen), daß $FIB(1) = 1$ und $FIB(2) = 1$ ist. Wenn wir nun zeigen können, daß $FIB(n) = FIB(n-1) + FIB(n-2)$ gilt, so haben wir damit alles bewiesen (insbesondere auch, daß für alle $n \in \mathbb{N}$ der Term $FIB(n)$ eine natürliche Zahl ergibt). Wir erkennen (quadratische Gleichung), daß α bzw. β die Nullstellen der Funktion $f(x) = x^2 - x - 1$ sind und haben zu zeigen, daß

$$\frac{\alpha^n - \beta^n}{\sqrt{5}} = \frac{\alpha^{n-1} - \beta^{n-1}}{\sqrt{5}} + \frac{\alpha^{n-2} - \beta^{n-2}}{\sqrt{5}}$$

ist. Diese Gleichung ist äquivalent mit

$$\alpha^n - \alpha^{n-1} - \alpha^{n-2} = \beta^n - \beta^{n-1} - \beta^{n-2}$$

bzw. mit

$$\alpha^{n-2}(\alpha^2 - \alpha - 1) = \beta^{n-2}(\beta^2 - \beta - 1) \ ,$$

eine Gleichung, die wegen $\alpha^2 - \alpha - 1 = 0$ und $\beta^2 - \beta - 1 = 0$ (α, β sind Nullstellen der angegebenen Funktion!) erfüllt ist. Hiemit ist alles gezeigt!

Beweis 2: („Linearisierung“ der Potenzen von α und β.) Da α und β die beiden Lösungen von $x^2 - x - 1 = 0$ sind, gilt für $x = \alpha$ und $x = \beta$ die Beziehung (x^2 kann durch $1 \cdot x + 1$ *linearisiert* werden):

$$x^2 = 1 \cdot x + 1 \qquad (14)$$

Mulitplizieren wir (14) mit x, so erhalten wir

$$x^3 = x^2 + x \qquad (15)$$

und mit Hilfe von (14)

$$x^3 = 2 \cdot x + 1 \ ,$$

so daß wir auch die 3. Potenz von α und β linearisiert haben. Nun fahren wir analog fort: durch Multiplizieren von (14) mit x^n erhalten wir $x^{n+2} = x^{n+1} + x^n$, so daß die Linearisierung von x^{n+2} durch jene von x^{n+1} und x^n einfach gewonnen werden kann. Wir erhalten der Reihe nach

$$\begin{array}{ccccc} x^3 & = & x^2 + x & = & 2 \cdot x + 1 \\ x^4 & = & x^3 + x^2 & = & 3 \cdot x + 2 \\ x^5 & = & x^4 + x^3 & = & 5 \cdot x + 3 \\ x^6 & = & x^5 + x^4 & = & 8 \cdot x + 5 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{array}$$

Nun *sieht* man für $x = \alpha$ und $x = \beta$ (auch ohne strenge vollständige Induktion):

$$\boxed{x^n = f_n \cdot x + f_{n-1}} \ .$$

Setzen wir hier α und β explizit ein, so erhalten wir

$$\begin{array}{ccc} \alpha^n & = & f_n \cdot \alpha + f_{n-1} \\ \beta^n & = & f_n \cdot \beta + f_{n-1} \ , \end{array}$$

woraus durch Subtraktion der beiden Gleichungen und wegen $\alpha - \beta = \sqrt{5}$ sofort folgt: $\boxed{f_n = \frac{\alpha^n - \beta^n}{\sqrt{5}}}$.

Nun können wir an den Beweis von

$$\sum_{n=2}^{\infty} n \, \frac{f_{n-1}}{2^n} = \sum_{n=2}^{\infty} n \, \frac{FIB(n-1)}{2^n} = 6$$

gehen, indem wir einsetzen:

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

Mit den schon angegebenen Abkürzungen α und β erhalten wir:

$$\sum_{n=2}^{\infty} n \, \frac{f_{n-1}}{2^n} = \sum_{n=2}^{\infty} \frac{n}{2^n} \, \frac{1}{\sqrt{5}} \, (\alpha^{n-1} - \beta^{n-1}) =$$

$$= \frac{1}{2\sqrt{5}} \left[\sum_{n=2}^{\infty} n \left(\frac{\alpha}{2}\right)^{n-1} - \sum_{n=2}^{\infty} n \left(\frac{\beta}{2}\right)^{n-1} \right] \ .$$

Da sowohl $\left|\frac{\alpha}{2}\right| < 1$ als auch $\left|\frac{\beta}{2}\right| < 1$ ist, können wir für beide Summen die Identität (3) anwenden (jeweils „$\cdots - 1$“, weil die Summation erst bei 2 beginnt) und erhalten nach kurzer Rechnung

$$= \frac{1}{2\sqrt{5}} \left[\left(\frac{1}{\left(1-\frac{\alpha}{2}\right)^2} - 1 \right) - \left(\frac{1}{\left(1-\frac{\beta}{2}\right)^2} - 1 \right) \right]$$

$$= \cdots = \boxed{6}\,.$$

Bemerkung: Einer Anregung von Herrn H. Haake verdanken wir folgende Möglichkeit, den Rückgriff auf die explizite Darstellung der Fibonacci-Zahlen zu umgehen (durch sehr „unbefangenen“ Umgang mit Reihen):

Durch Ausmultiplizieren kann die Identität

$$(f_1 + f_2 x + f_3 x^2 + \ldots)(1 - x - x^2) = 1$$

leicht bestätigt werden. Die Multiplikation mit x^2 führt zu

$$\sum_{n=2}^{\infty} f_{n-1} x^n = \frac{x^2}{1 - x - x^2} \quad \text{für} \quad |x| < \frac{\sqrt{5}-1}{2}\,.$$

Der Gültigkeitsbereich („Konvergenzradius“) $|x| < \frac{\sqrt{5}-1}{2}$ ist auch ohne weitere exakte Betrachtungen zumindest plausibel, da in diesem keine Nullstelle des Nenners (Polstelle) liegt: $\frac{\sqrt{5}-1}{2}$ ist die betraglich kleinere der beiden Polstellen $\frac{-1\pm\sqrt{5}}{2}$. (Gliedweises) Differenzieren führt zu

$$\sum_{n=2}^{\infty} n f_{n-1} x^{n-1} = \frac{x(2-x)}{(1-x-x^2)^2}$$

für $|x| < \frac{\sqrt{5}-1}{2}$. Einsetzen von $x = \frac{1}{2}$ (liegt im erlaubten Bereich!) ergibt

$$\sum_{n=2}^{\infty} n \frac{f_{n-1}}{2^{n-1}} = 12\,,$$

also ist

$$E(KK) = \sum_{n=2}^{\infty} n \frac{f_{n-1}}{2^n} = 6\,.$$

Literatur

Engel, A. (1973, 1976): Wahrscheinlichkeitsrechnung und Statistik (2 Bände). Klett, Stuttgart.

Gardner, M. (1974): On the Paradoxical Situations That Arise From Nontransitive Relations. In: *Scientific American*, 231 (Oktober 1974), 120–125.

Humenberger, H. (2000): Überraschendes bei Münzwurfserien. In: *Stochastik in der Schule* **20**, 1, 4–20.

Li, S.-Y.R. (1980): A Martingale Approach to the Study of Occurence of Sequence Patterns in Repeated Experiments. In: *The Annals of Probability* **8**, 6, 1171–1176.

Reichel, H.-C.; Hanisch, G.; Müller, R. ([3]1992): Wahrscheinlichkeitsrechnung und Statistik. Mathematik für Schule und Praxis (Hrsg.: H.-C. Reichel), Band 1. Hölder-Pichler-Tempsky, Wien.

Szekely, G.J. (1990): Paradoxa – klassische und neue Überraschungen aus Wahrscheinlichkeitsrechnung und mathematischer Statistik. Harri Deutsch, Frankfurt.

Waterman, M.S. (1995): Introduction to Computational Biology. Chapman and Hall.

Adresse des Autors

Hans Humenberger
IEEM, Fachbereich Mathematik
Universität Dortmund

D-44 221 Dortmund

e-mail:
hans.humenberger@math.uni-dortmund.de