Zeitschrift: Elemente der Mathematik
Herausgeber: Schweizerische Mathematische Gesellschaft
Band: 55 (2000)

Artikel: Die Thue-Morse-Folge
Autor: Puchta, Jan-Christoph / Spilker, Jürgen
DOI: https://doi.org/10.5169/seals-5634

Nutzungsbedingungen
Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation
L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use
The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 07.01.2026

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

Elem. Math. 55 (2000) 110 – 122
0013-6018/00/030110-13 $ 1.50+0.20/0

© Birkhäuser Verlag, Basel, 2000

Elemente der Mathematik

Die Thue-Morse-Folge

Jan-Christoph Puchta
Jürgen Spilker

Jan-Christoph Puchta wurde 1973 in Eckernförde geboren. Er studierte Mathematik in Freiburg und promovierte bei D. Wolke über Fragen der Primzahlverteilung. Seine Forschungsgebiete sind elementare und analytische Zahlentheorie.

Jürgen Spilker wurde 1935 in Berlin geboren. Er hat Mathematik und Physik in Göttingen studiert und promovierte dort bei H. Grauert. Seit 30 Jahren lehrt er an der Universität Freiburg (Breisgau) und engagiert sich besonders bei der Ausbildung der Lehramtskandidaten. Sein Forschungsgebiet war zunächst die Theorie automorpher Formen, heute ist es die Zahlentheorie, insbesondere das Gebiet zahlentheoretische Funktionen mit Beweismethoden aus der reellen Analysis. Sein Hobby ist die Postgeschichte Deutschlands 1945–1950. Er sammelt Briefe und postalische Belege aus dieser Zeit.

0 Einleitung

Jede natürliche Zahl $n \in \mathbb{N}_0$ hat eine eindeutige Dualdarstellung

$$n = \varepsilon_0 + \varepsilon_1 2 + \varepsilon_2 2^2 + \cdots + \varepsilon_N 2^N$$

Zum Testen von Computerprogrammen, für numerische Algorithmen und in der Statistik benötigt man oft Zufallszahlen. Da man echte Zufallszahlen nicht ökonomisch erhalten kann, benutzt man in der Praxis oft Zahlenfolgen, die einerseits schnell berechenbar sind, andererseits gewünschte Eigenschaften von Zufallsfolgen besitzen. Die Thue-Morse-Folge

$$1, -1, -1, 1, -1, 1, 1, -1, -1, 1, 1, -1, 1, -1, \ldots$$

ist ein Beispiel für eine derartige Folge. In der vorliegenden Arbeit wird gezeigt, dass die Thue-Morse-Folge sehr zufällig erscheint, so lange nur einzelne Glieder betrachtet werden (so treten etwa ungefähr gleich viele 1 und -1 auf), andererseits dass Regelmässigkeiten auftreten, sobald man mehrere Glieder betrachtet (so folgt auf eine 1 öfter eine -1 als eine weitere 1). Die Beweise basieren auf einem Lemma von Gelfond. Dieses wird auch auf ein anderes zahlentheoretisches Problem angewendet. So lässt sich damit die Vermutung von S. Ulam beweisen, dass die Wahrscheinlichkeit, dass eine natürliche Zahl und ihre Quersumme beide durch 13 teilbar sind, gleich $1/169$ ist. *jk*

mit $\varepsilon_k \in \{0,1\}$. Sei $s(n) := \sum_{0 \le k \le N} \varepsilon_k$ die Quersumme. Dann heißt

$$f(n) := (-1)^{s(n)}, \quad n \in \mathbb{N}_0$$

die Thue-Morse-Folge. In dieser Arbeit wird die Frage untersucht, wie zufällig ihre Werte $1, -1$ verteilt sind. Zunächst wird bewiesen, daß ihre beiden Werte gleich oft auftreten, genauer, daß der Mittelwert

$$M(f) := \lim_{x \to \infty} \frac{1}{x} \sum_{0 \le n < x} f(n)$$

existiert und den Wert 0 hat. Das gilt auch, wenn man nur über eine Restklasse l mod m summiert (Satz 3). Eine zweite Möglichkeit, den Grad der Zufälligkeit zu messen, besteht darin, die Wohlverteilung der Folgen $n \longmapsto f(n)\, f(n+b)$ mit $b \in \mathbb{N}$ zu zeigen. Sie haben ebenfalls Mittelwerte $k(b)$. Eine explizite Formel der "Korrelations-Koeffizienten" $k(b)$ ist nicht bekannt, aber sie lassen sich rekursiv berechnen (Satz 4), und ihr Mittelwert existiert und ist 0 (Satz 10). Das gilt auch dann noch, wenn man die Mittelung nur über eine Restklasse l mod m erstreckt (Satz 13).

Das wichtigste Hilfsmittel der Beweise ist das Lemma 1 von Gelfond. Es wird in Abschnitt 1 bewiesen. Bevor damit die Thue-Morse-Folge untersucht wird, wenden wir es in Abschnitt 2 auf ein interessantes zahlentheoretisches Problem an.

1 Das Lemma von Gelfond

Wir beweisen die Abschätzung einer Exponentialsumme, welche auf A.O. Gelfond zurückgeht.

Lemma 1 von Gelfond ([3], S. 261) *Sei $g \ge 2$ eine ganze Zahl und $s(n)$ die Quersumme der g-adischen Darstellung der natürlichen Zahl n. Dann gilt für alle $\alpha \in \mathbb{R}, \xi = \frac{j}{m} \in \mathbb{Q}$ mit $0 < j < m, (m, g-1) = 1$:*

$$\lim_{x \to \infty} \frac{1}{x} \sum_{0 \le n < x} e^{2\pi i(\alpha n + \xi s(n))} = 0.$$

Beweis für $g = 2$ ([2], S. 582). Sei $\alpha \in \mathbb{R}, \ \xi \in \mathbb{R} \setminus \mathbb{Z}$,

$$T(x) := \sum_{0 \le n < x} e^{2\pi i(\alpha n + \xi s(n))}$$

und zunächst $x = 2^N$. Jede ganze Zahl $0 \le n < 2^N$ hat genau eine Darstellung

$$n = \varepsilon_0 + \varepsilon_1\, 2 + \varepsilon_2\, 2^2 + \ldots + \varepsilon_{N-1}\, 2^{N-1} \quad \text{mit} \quad \varepsilon_j \in \{0,1\}.$$

Hieraus folgt die grundlegende Identität

$$\sum_{0 \le n < 2^N} e^{2\pi i(\alpha n + \xi s(n))} = \prod_{0 \le j < N} (e^0 + e^{2\pi i(\alpha 2^j + \xi)})$$

sowie

$$|T(2^N)| = 2^N \prod_{0 \le j < N} |\cos(\pi(\alpha 2^j + \xi))|.$$

In dem Produkt faßt man die Faktoren – solange es geht – paarweise zusammen:

$$|T(2^N)| \le c2^N \Big[\max_{t \in \{1, 2^2, 2^4, \ldots\}} |\cos \pi(\alpha t + \xi) \cos \pi(2\alpha t + \xi)| \Big]^{\frac{N}{2}}.$$

Sei $\|y\| := \min\{|y - n| : n \in \mathbb{Z}\}, y \in \mathbb{R}$. Weil ξ keine ganze Zahl ist, gibt es ein $\delta > 0$ mit $\|y + 2\xi\| \ge \delta$ oder $\|y + \xi\| \ge \delta$ für alle reellen y. Somit gilt $\|\alpha t + \xi\| \ge \frac{\delta}{2}$ oder $\|2\alpha t + \xi\| \ge \delta$ für alle $t \in \mathbb{R}$. Folglich existieren Konstanten $a = a(\xi) < 1$ und $b = b(\xi) < 1$ mit

(1) $|T(2^N)| \le c2^N \, a^{\frac{N}{2}} = c2^{Nb}, \quad \forall\, N \in \mathbb{N},$

folglich ist $\lim\limits_{N \to \infty} \frac{1}{2^N} T(2^N) = 0$. Im allgemeinen Fall $x \in \mathbb{R}, \ge 1$ setzt man

$$[x] = 2^{n_1} + 2^{n_2} + \ldots + 2^{n_k} \quad \text{mit} \quad n_1 > n_2 > \ldots > n_k \ge 0$$

und erhält

$$T(x) = T(2^{n_1}) + T(2^{n_2})\, e^{2\pi i(\alpha 2^{n_1} + \xi)} + T(2^{n_3})\, e^{2\pi i(\alpha(2^{n_1} + 2^{n_2}) + 2\xi)} + \ldots$$

Aus der Abschätzung (1) des Spezialfalles folgt

$$\begin{aligned} |T(x)| &\le c\, (2^{n_1 b} + 2^{n_2 b} + \ldots + 2^{n_k b}) \\ &\le c \sum_{0 \le n \le n_1} 2^{nb} \\ &< c'\, 2^{n_1 b} \qquad \text{mit } c' := c \frac{2^b}{2^b - 1} \\ &\le c'\, x^b \end{aligned}$$

und damit wegen $b < 1$ endlich $\lim\limits_{x \to \infty} \frac{1}{x}\, T(x) = 0$. □

2 Eine zahlentheoretische Anwendung des Lemmas von Gelfond

S. Ulam hat gefragt [1], ob die Wahrscheinlichkeit, daß eine natürliche Zahl durch 13 teilbar und auch ihre Quersumme in der Dezimaldarstellung durch 13 teilbar ist, den erwarteten Wert $\frac{1}{169}$ hat. Das ist der Fall, denn es gilt der

Satz 2 ([3], S. 263) *Sei $g \ge 2$ eine ganze Zahl und $s(n)$ die Quersumme der g-adischen Darstellung von der natürlichen Zahl n. Dann gilt für alle $a, b \in \mathbb{Z}$, $l, m \in \mathbb{N}$ mit $(m, g-1) = 1$:*

$$A_x := \frac{1}{x} \#\{0 \le n < x : n \equiv a \mod l,\ s(n) \equiv b \mod m\} \xrightarrow[x \to \infty]{} \frac{1}{lm}.$$

Beispiel von Ulam: $g = 10,\ a = b = 0,\ l = m = 13$.

Beweis. Die geometrische Reihe läßt sich summieren:

$$(2)\quad \sum_{0\le j<m} e^{2\pi i\frac{n-l}{m}j} = \begin{cases} m & \text{falls} \quad n\equiv l \mod m \\ 0 & \text{falls} \quad n\not\equiv l \mod m. \end{cases}$$

Hiermit folgt

$$\begin{aligned} A_x &= \frac{1}{x}\sum_{0\le n<x}\Big(\frac{1}{l}\sum_{0\le k<l} e^{2\pi i\frac{n-a}{l}k}\,\frac{1}{m}\sum_{0\le j<m} e^{2\pi i\frac{s(n)-b}{m}j}\Big) \\ &= \frac{1}{lm}\sum_{0\le k<l} e^{-2\pi i\frac{a}{l}k}\sum_{0\le j<m} e^{-2\pi i\frac{b}{m}j}\Big(\frac{1}{x}\sum_{0\le n<x} e^{2\pi i(\frac{k}{l}n+\frac{j}{m}s(n))}\Big). \end{aligned}$$

Für $j\neq 0$ strebt wegen Lemma 1 von Gelfond der innere Ausdruck bei $x\to\infty$ gegen 0, also konvergiert A_x für $x\to\infty$ gegen

$$\frac{1}{lm}\sum_{0\le k<l} e^{-2\pi i\frac{a}{l}k}\Big(\lim_{x\to\infty}\frac{1}{x}\sum_{0\le n<x} e^{2\pi i\frac{k}{l}n}\Big).$$

Der Grenzwert in der Klammer ist 1 für $k=0$ und sonst 0. Damit ist bewiesen, daß A_x gegen $\frac{1}{lm}$ konvergiert. □

3 Restklassen-Verteilung der Thue-Morse-Folge

Wir verwenden nun das Gelfond-Lemma, um die Thue-Morse-Folge $f(n)$ zu studieren. Sie verhält sich im Kleinen zufällig und ihre Werte 1 und -1 sind wohlverteilt, ja sogar in jeder festen Restklasse wohlverteilt. Um das zu präzisieren, definieren wir

$$M_{l,m}(f) := \lim_{x\to\infty}\frac{1}{x}\sum_{\substack{0\le n<x\\ n\equiv l \mod m}} f(n)$$

und zeigen den

Satz 3 *Für alle ganzen l,m mit $m\ge 1$ gilt $M_{l,m}(f)=0$.*

Beweis. Mit (2) folgt

$$\begin{aligned} \frac{1}{x}\sum_{\substack{0\le n<x\\ n\equiv l \mod m}} f(n) &= \frac{1}{x}\sum_{0\le n<x}\Big(\frac{1}{m}\sum_{0\le j<m} e^{2\pi i\frac{n-l}{m}j} f(n)\Big) \\ &= \frac{1}{m}\sum_{0\le j<m} e^{-2\pi i\frac{l}{m}j}\Big(\frac{1}{x}\sum_{0\le n<x} e^{2\pi i(\frac{j}{m}n+\frac{1}{2}s(n))}\Big). \end{aligned}$$

Läßt man in dieser Darstellung $x\to\infty$ gehen, dann verschwinden die Ausdrücke in den Klammern (Lemma 1 für $g=2$). Also konvergiert auch die linke Seite gegen 0. □

4 Mittelwerte von Produkten geshifteter Thue-Morse-Folgen

Wir betrachten jetzt geshiftete Thue-Morse-Folgen

$$f_a(n) := f(n+a), \qquad a \in \mathbb{N}_0$$

sowie Produkte von zwei solchen

$$f_{a,b}(n) := f(n+a)\, f(n+b),\ a, b \in \mathbb{N}_0,\ 0 \le a < b.$$

Wegen

$$\sum_{0 \le n < x} f(n+a)\, f(n+b) = \sum_{a \le m < x+a} f(m)\, f(m+b-a)$$

gilt $M(f_{a,b}) = M(f_{0,b-a})$, sofern die Mittelwerte existieren. Man kann deshalb bei den folgenden Mittelwertuntersuchungen $a = 0$ setzen. Wir zeigen zunächst, daß jede Folge $f_{0,b}$ einen Mittelwert $k(b)$ hat und die $k(b)$ eine einfache Rekursionsformel erfüllen.

Satz 4 *Die "Korrelations-Koeffizienten" $k(b) := M(f_{0,b}), b \in \mathbb{N}_0$ der Thue-Morse-Folge existieren, und es gilt*

$$\begin{aligned} k(0) &= 1 \\ k(2b) &= k(b), && b \in \mathbb{N}_0 \\ k(2b+1) &= -\frac{1}{2}k(b) - \frac{1}{2}k(b+1), && b \in \mathbb{N}_0. \end{aligned}$$

Hieraus berechnet man sofort:

b	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18 ...
$k(b)$	1	$-\frac{1}{3}$	$-\frac{1}{3}$	$\frac{1}{3}$	$-\frac{1}{3}$	0	$\frac{1}{3}$	0	$-\frac{1}{3}$	$\frac{1}{6}$	0	$-\frac{1}{6}$	$\frac{1}{3}$	$-\frac{1}{6}$	0	$\frac{1}{6}$	$-\frac{1}{3}$	$\frac{1}{12}$	$\frac{1}{6}$ ...

Zunächst ein

Lemma 5 *Zu jeder natürlichen Zahl l existiert eine Folge $g_l : \mathbb{N}_0 \to \{1, 0, -1\}$, welche die Periode 2^l hat und*

$$\limsup_{x \to \infty} \frac{1}{x} \sum_{0 \le n < x} \left| f_{0,1}(n) - g_l(n) \right| \le \frac{1}{2^l}$$

erfüllt.

Beweis zu Lemma 5. Wir definieren eine Folge $g_l : \mathbb{N}_0 \to \{1, 0, -1\}$ durch

$$\begin{aligned} g_l(n) &= f_{0,1}(n), \quad 0 \le n < 2^l - 1 \\ g_l(2^l - 1) &= 0, \\ g_l(n + 2^l) &= g_l(n), \qquad \forall\, n \in \mathbb{N}_0. \end{aligned}$$

Sie hat die Periode 2^l. Ferner gilt

(3) $g_l(n) = f_{0,1}(n), \quad \forall\, n \not\equiv 2^l - 1 \bmod 2^l,$

denn jedes solche $n = \varepsilon_0 + \varepsilon_1\, 2 + \varepsilon_2\, 2^2 + \ldots$ hat unter seinen Ziffern $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{l-1}$ mindestens eine 0. Ist $\varepsilon_{k-1} = 0$ und der Index minimal gewählt, dann hat n die Dualdarstellung

$$n = \ldots \varepsilon_{k+1}\ \varepsilon_k\ 0\ 1\ 1\ \ldots 1 \quad \text{mit} \quad 1 \le k \le l,$$

und

$$n + 1 = \ldots \varepsilon_{k+1}\ \varepsilon_k\ 1\ 0\ 0 \ldots 0$$

und

$$f_{0,1}(n) = (-1)^k$$

und also

$$\begin{aligned} g_l(n) &= g_l(\varepsilon_{l-1}\ \varepsilon_{l-2} \ldots \varepsilon_k\ 0\ 1\ 1 \ldots 1) \\ &= f_{0,1}(\varepsilon_{l-1}\ \varepsilon_{l-2} \ldots \varepsilon_k\ 0\ 1\ 1 \ldots 1) \\ &= f_{0,1}(n). \end{aligned}$$

Damit ist (3) gezeigt. Hieraus folgt die Behauptung

$$\limsup_{x\to\infty} \frac{1}{x} \sum_{0 \le n < x} |f_{0,1}(n) - g_l(n)| \le \limsup_{x\to\infty} \frac{1}{x} \sum_{\substack{0 \le n < x \\ n \equiv 2^l - 1 \bmod 2^l}} 1 \le \lim_{x\to\infty} \frac{1}{x}\, \frac{x+1}{2^l} = \frac{1}{2^l}. \square$$

Beweis von Satz 4. Hängt man an die Dualdarstellung einer natürlichen Zahl n eine 0 an, dann erkennt man

(4) $f(2n) = f(n), \quad \forall\, n \in \mathbb{N}_0.$

Ebenso sieht man durch Anhängen von 1

(5) $f(2n + 1) = -f(n), \quad \forall\, n \in \mathbb{N}_0.$

1) Wegen $f(n)^2 = 1$ ist $k(0) = 1$. Wir beweisen jetzt $k(1) = -\frac{1}{3}$. Mit (4), (5) ist für jedes natürliche n

$$\begin{aligned} f_{0,1}(2n) &= -f(n)^2 = -1, \\ f_{0,1}(2n + 1) &= -f(n)\, f(n + 1) = -f_{0,1}(n). \end{aligned}$$

Hieraus folgt

$$\begin{aligned} S_l &:= \sum_{0\le n<2^l} f_{0,1}(n) \\ &= \sum_{0\le n<2^{l-1}} f_{0,1}\,(2n) + \sum_{0\le n<2^{l-1}} f_{0,1}\,(2n+1) \\ &= -2^{l-1} - S_{l-1}, \quad \forall\, l \in \mathbb{N}. \end{aligned}$$

Durch Iteration ergibt sich

$$S_l = -2^{l-1} + 2^{l-2} - 2^{l-3} \pm \ldots + (-1)^{l-1}\, 2 + (-1)^{l-1}\, S_1 = -\frac{2^l}{3} - (-1)^l\, \frac{2}{3},$$

da $S_1 = 0$ ist. Die Folge g_l aus Lemma 5 hat also den Mittelwert $\frac{1}{2^l} S_l = -\frac{1}{3} - (-1)^l\, \frac{1}{2^{l-1}3}$. Läßt man in der Ungleichung

$$\limsup_{x\to\infty} \frac{1}{x} \sum_{0\le n<x} f_{0,1}(n) \le \frac{1}{2^l}\, S_l + \frac{1}{2^l}$$

$l \to \infty$ gehen, dann erhält man

$$\limsup_{x\to\infty} \frac{1}{x} \sum_{0\le n<x} f_{0,1}(n) \le -\frac{1}{3}.$$

In analoger Weise erkennt man

$$\liminf_{x\to\infty} \frac{1}{x} \sum_{0\le n<x} f_{0,1}(n) \ge -\frac{1}{3}.$$

Wegen der letzten beiden Ungleichungen existiert der Grenzwert $\lim\limits_{x\to\infty} \frac{1}{x} \sum\limits_{0\le n<x} f_{0,1}(n)$, und er hat den Wert $-\frac{1}{3}$. Damit ist der Korrelations-Koeffizient $k(1)$ berechnet.

2) Berechnung von $k(b)$ für $b > 1$. Mit (4), (5) findet man für alle natürlichen b

$$\begin{aligned} \sum_{0\le n<2^l} f_{0,2b}(n) &= \sum_{0\le 2m<2^l} f(2m)\, f(2m+2b) + \sum_{0\le 2m+1<2^l} f(2m+1)\, f(2m+2b+1) \\ &= \sum_{0\le m<2^{l-1}} f(m)\, f(m+b) \quad + \sum_{0\le m<2^{l-1}} f(m)\, f(m+b) \\ &= 2 \sum_{0\le m<2^{l-1}} f_{0,b}(m). \end{aligned}$$

Aus der Identität

$$\frac{1}{2^l} \sum_{0\le n<2^l} f_{0,2b}(n) = \frac{1}{2^{l-1}} \sum_{0\le n<2^{l-1}} f_{0,b}(n)$$

folgt: wenn $k(b)$ existiert, dann existiert auch $k(2b)$, und es gilt $k(2b) = k(b)$. In analoger Weise folgt

$$\frac{1}{2^l} \sum_{0 \le n < 2^l} f_{0,2b+1}(n) = -\frac{1}{2^l} \sum_{0 \le n < 2^{l-1}} f_{0,b}(n) - \frac{1}{2^l} \sum_{0 \le n < 2^{l-1}} f_{0,b+1}(n)$$

und hieraus: wenn $k(b)$ und $k(b+1)$ existieren, dann existiert auch $k(2b+1)$, und es gilt $k(2b+1) = -\frac{1}{2}\ k(b) - \frac{1}{2}\ k(b+1)$. Durch vollständige Induktion über b folgt aus 1) und 2) die Behauptung. □

Die Rekursionsformeln in Satz 4 haben die

Folgerung 6 *Für alle $b \in \mathbb{N}$ ist $|k(b)| \le \frac{1}{3}$.*

Eine explizite Formel für die Korrelations-Koeffizienten $k(b)$ ist nicht bekannt. Es gibt jedoch eine Matrizendarstellung, nämlich

Satz 7 *Sei $A := \left(\begin{smallmatrix} 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} \end{smallmatrix}\right)$ und $I := \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$.*
Dann gilt für jedes $b = \varepsilon_0 + \varepsilon_1 2 + \ldots + \varepsilon_{k-1}\ 2^{k-1} + 2^k,\ k \ge 1$:

$$\begin{pmatrix} k(b) \\ k(b+1) \end{pmatrix} = -\frac{1}{3}\ I^{\varepsilon_0}\ A\ I^{\varepsilon_0+\varepsilon_1}\ A\ I^{\varepsilon_1+\varepsilon_2}\ A \ldots A\ I^{\varepsilon_{k-2}+\varepsilon_{k-1}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Beweis. Wenn $\varepsilon_0 = 0$, also b gerade ist, dann gilt nach Satz 4

$$\begin{pmatrix} k(b) \\ k(b+1) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} k(\frac{b}{2}) \\ k(\frac{b}{2}+1) \end{pmatrix} = A \begin{pmatrix} k(\frac{b}{2}) \\ k(\frac{b}{2}+1) \end{pmatrix}.$$

Wenn $\varepsilon_0 = 1$, also b ungerade ist, dann ist

$$\begin{pmatrix} k(b) \\ k(b+1) \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & -\frac{1}{2} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} k(\frac{b-1}{2}) \\ k(\frac{b-1}{2}+1) \end{pmatrix} = I\ A\ I \begin{pmatrix} k(\frac{b-1}{2}) \\ k(\frac{b+1}{2}) \end{pmatrix}.$$

In beiden Fällen gilt also

$$\begin{pmatrix} k(b) \\ k(b+1) \end{pmatrix} = I^{\varepsilon_0} A\ I^{\varepsilon_0} \begin{pmatrix} k([\frac{b}{2}]) \\ k([\frac{b}{2}]+1) \end{pmatrix}.$$

Wegen $[\frac{b}{2}] = \varepsilon_1 + \varepsilon_2\ 2 + \ldots + \varepsilon_{k-1}\ 2^{k-2} + 2^{k-1}$ kann man diese Identität iterieren und erhält

$$\begin{pmatrix} k(b) \\ k(b+1) \end{pmatrix} = I^{\varepsilon_0}\ A\ I^{\varepsilon_0}\ \cdot I^{\varepsilon_1}\ A\ I^{\varepsilon_1} \ldots I^{\varepsilon_{k-1}} A\ I^{\varepsilon_{k-1}} \begin{pmatrix} k(1) \\ k(2) \end{pmatrix}.$$

Mit $k(1) = k(2) = -\frac{1}{3}$ und $A\ I^{\varepsilon_{k-1}} \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ folgt die Behauptung. □

Bemerkung Die Potenzen der Matrizen I und A lassen sich leicht berechnen: Sei $\varepsilon \in \mathbb{N}$ und E die zweireihige Einheitsmatrix, dann gilt

$$I^\varepsilon = \begin{cases} E & \varepsilon \text{ gerade} \\ I & \varepsilon \text{ ungerade} \end{cases}, \quad A^\varepsilon = \begin{pmatrix} 1 & 0 \\ -\frac{1}{3}(1-(-\frac{1}{2})^\varepsilon) & (-\frac{1}{2})^\varepsilon \end{pmatrix}. \qquad \square$$

Bisher haben wir Produkte von zwei geshifteten Thue-Morse-Folgen betrachtet. Man wird erwarten, daß auch Produkte von drei oder mehr Shiftungen einen Mittelwert haben. Das ist wahr, und diese Mittelwerte verschwinden immer.

Satz 8 *Seien $0 \le a < b < c$ ganze Zahlen. Dann hat die Funktion*

$$f_{a,b,c}(n) := f(n+a)\, f(n+b)\, f(n+c)$$

den Mittelwert 0.

Beweis für den Fall $a = 0$, $b = 1$ (der allgemeine Fall läßt sich ähnlich behandeln). Sei $l \in \mathbb{N}$ und g_l nach Lemma 5 gewählt. Dann ist die 2^l-periodische Funktion g_l Linearkombination von Exponentialfunktionen, etwa

$$g_l(n) = \sum_{0 \le k < 2^l} \gamma_k\, e^{2\pi i \frac{k}{2^l} n}, \quad \gamma_k \in \mathbb{C}.$$

Damit folgt

$$\begin{aligned} \Big|\frac{1}{x} \sum_{0 \le n < x} f_{0,1,c}(n)\Big| &\le \Big|\frac{1}{x} \sum_{0 \le n < x} (f_{0,1}(n) - g_l(n)) f(n+c)\Big| + \Big|\frac{1}{x} \sum_{0 \le n < x} g_l(n) f(n+c)\Big| \\ &\le \frac{1}{x} \sum_{0 \le n < x} |f_{0,1}(n) - g_l(n)| + \sum_{0 \le k < 2^l} \Big(|\gamma_k| \, \Big|\frac{1}{x} \sum_{0 \le n < x} e^{2\pi i \frac{k}{2^l} n} f(n+c)\Big|\Big). \end{aligned}$$

Der Ausdruck im letzten Betrag ist $\Big|\frac{1}{x} \sum_{c \le m < x+c} e^{2\pi i (\frac{k}{2^l} m + \frac{1}{2} s(m))}\Big|$. Er strebt gegen 0 bei $x \to \infty$ (Lemma 1, $g = 2$). Also ist

$$\limsup_{x \to \infty} \Big|\frac{1}{x} \sum_{0 \le n < x} f_{0,1,c}(n)\Big| \le \frac{1}{2^l}, \quad \forall\, l \in \mathbb{N},$$

und es existiert $M_{0,1,c} = 0$ für jedes $c > 1$. $\square$

Bemerkung Auch das Produkt von k geshifteten Funktionen $\prod_{1 \le j \le k} f(n + a_j)$ mit $0 \le a_1 < a_2 < \ldots < a_k$ hat einen Mittelwert. Er verschwindet, wenn k eine ungerade Zahl ist.

5 Restklassen-Verteilung der Korrelations-Koeffizienten

Die interessantesten Daten der Thue-Morse-Folge sind die Korrelations-Koeffizienten $k(b)$, $b \in \mathbb{N}_0$. In diesem Abschnitt wird gezeigt, daß sie auf jeder Restklasse wohlverteilt sind (Satz 13). Zunächst beweisen wir eine einfache Formel für $\sum\limits_{0 \le b < x} k(b)$:

Satz 9 *Für alle $b \in \mathbb{N}$ gilt*

$$\begin{aligned} k(0) + k(1) + \ldots + k(2b-1) &= \frac{1}{2}(1 - k(b)) \\ k(0) + k(1) + \ldots + k(2b) &= \frac{1}{2}(1 + k(b)). \end{aligned}$$

Beweis durch Induktion über b. Für $b = 1$ ist

$$\begin{aligned} k(0) + k(1) &= \frac{2}{3} = \frac{1}{2}(1 - k(1)) \\ k(0) + k(1) + k(2) &= \frac{1}{3} = \frac{1}{2}(1 + k(1)). \end{aligned}$$

Gilt die Behauptung für ein natürliches b, dann folgen mit den Rekursionsformeln aus Satz 4 die Gleichungen für $b + 1$:

$$\begin{aligned} k(0) + k(1) + \ldots + k(2b+1) &= \frac{1}{2}(1 + k(b)) \quad + k(2b+1) = \frac{1}{2}(1 - k(b+1)) \\ k(0) + k(1) + \ldots + k(2b+2) &= \frac{1}{2}(1 - k(b+1)) + k(2b+2) = \frac{1}{2}(1 + k(b+1)). \end{aligned}$$

□

Weil $|k(b)| \le 1$ für jedes $b \in \mathbb{N}_0$ ist, folgt $\left| \sum\limits_{0 \le b < x} k(b) \right| \le 1$, $\forall\, x \ge 1$ aus Satz 9 und somit der

Satz 10 *Die Korrelations-Funktion $k : b \mapsto k(b)$ hat den Mittelwert* 0.

Obwohl $k(2^l) = -\frac{1}{3}$ für alle natürlichen l gilt, konvergieren die Mittel der Funktionen $|k|$ und k^2 gegen 0. Es gilt nämlich der

Satz 11 *Für jedes reelle $x > 1$ ist*

1) $\sum\limits_{1 \le b < x} |k(b)| < x^{\frac{3}{4}}$,

2) $\sum\limits_{1 \le b < x} k^2(b) < x^{\frac{1}{2}}$.

Beweis. Wir zeigen zunächst 2) und damit anschließend 1).
Sei a eine natürliche Zahl und $k(a) =: x,\ k(a+1) =: y$. Aus den Rekursionsformeln in Satz 4 folgt

$$\begin{aligned} k(2a) &= x \\ k(2a+1) &= -\frac{1}{2}(x+y) \\ k(2a+2) &= y \end{aligned}$$

und

$$\begin{aligned} k(4a) &= x \\ k(4a+1) &= -k(4a+3) = -\frac{1}{4}(x-y) \\ k(4a+2) &= -\frac{1}{2}(x+y). \end{aligned}$$

Somit gilt

$$\sum_{4a \le b < 4(a+1)} k^2(b) = x^2 + \frac{1}{8}(x-y)^2 + \frac{1}{4}(x+y)^2 = \frac{1}{8}(11\ x^2 + 2xy + 3\ y^2).$$

Wegen $2xy \le x^2 + y^2$ ist

$$\sum_{4a \le b < 4(a+1)} k^2(b) \le \frac{1}{2}(3x^2 + y^2).$$

Sind M und N natürliche Zahlen mit $M < N$, dann ergibt sich hieraus die Abschätzung

$$\begin{aligned} \sum_{4M \le b < 4N} k^2(b) &= \sum_{M \le a < N} \ \sum_{4a \le b < 4(a+1)} k^2(b) \\ &\le \frac{1}{2} \sum_{M \le a < N} (3\ k^2(a) + k^2(a+1)) \\ &\le 2 \sum_{M \le a < N} k^2(a) + \frac{1}{2}\ k^2(N). \end{aligned}$$

Wir setzen jetzt $K(x) := \sum_{1 \le b < x} k^2(b)$. Wegen $|k(N)| \le \frac{1}{3}$ (Folgerung 6) folgt für alle $N \ge M$

$$\begin{aligned} K(4N) - K(4M) &\le 2(K(N) - K(M)) + \frac{1}{18} \\ K(4N) - 2K(N) &\le K(4M) - 2K(M) + \frac{1}{18}. \end{aligned}$$

Sei y reell, ≥ 1 und $N := \min\{n \in \mathbb{N} : n \ge y\}$. Dann ergibt sich

$$\begin{aligned} K(4y) - 2K(y) &\le K(4N) - 2K(N-1) \\ &\le K(4N) - 2K(N) + \frac{2}{9} \\ &\le K(4) - 2K(1) + \frac{5}{18} = \frac{11}{18}. \end{aligned}$$

Diese für alle reellen $y \geq 1$ gültige Abschätzung iterieren wir:

$$K(4^2 y) \leq 2K(4y) + \frac{11}{18} \leq 2^2 K(y) + 2\,\frac{11}{18} + \frac{11}{18}.$$

So erhält man für jedes natürliche l

$$K(4^l\, y) \leq 2^l\, K(y) + (2^{l-1} + 2^{l-2} + \ldots + 1)\,\frac{11}{18} \leq 2^l \Big(K(y) + \frac{11}{18}\Big).$$

Sei endlich $x > 1$ und $l \in \mathbb{N}_0$ bestimmt durch $4^l \leq x < 4^{l+1}$; ferner $4^l\, y := x$. Dann folgt

$$K(x) \leq 2^{\frac{\log x}{\log 4}} \Big(K(4) + \frac{11}{18}\Big).$$

Wegen $2^{\log x} = x^{\log 2}$ und $K(4) = \frac{1}{3}$ ist das die Behauptung 2).

Aus 2) folgt mit der Ungleichung von Cauchy und Schwarz leicht 1), denn ist $x > 1$ gegeben, dann ist

$$\Big(\sum_{1 \leq b < x} |k(b)|\Big)^2 \leq \sum_{1 \leq b < x} 1 \cdot \sum_{1 \leq b < x} k^2(b) < x^{\frac{3}{2}}.$$

□

Bemerkung Man kann mit geeigneten Konstanten $c_2 > c_1 > 0$ sogar

$$c_1 x^{0,34} < \sum_{1 \leq b < x} k^2(b) < c_2\, x^{0,39}$$

für alle $x > 1$ zeigen. Das genaue Wachstum der Funktion $K(x)$ kennen wir nicht.

Aus Satz 11 ergibt sich sofort der

Satz 12 $M(|k|) = 0$.

Hieraus erkennt man schließlich, daß sich die Korrelations-Koeffizienten $k(b)$, $b \in \mathbb{N}_0$ auf die Restklassen wohlverteilen:

Satz 13 *Für alle ganzen Zahlen l, m mit $m > 0$ gilt $M_{l,m}(k) = 0$.*

Beweis. Sind l und $m > 0$ gegeben, dann folgt für jedes reelle $x > 1$

$$\frac{1}{x}\Big|\sum_{\substack{0 \leq b < x \\ b \equiv l \bmod m}} k(b)\Big| \leq \frac{1}{x} \sum_{\substack{0 \leq b < x \\ b \equiv l \bmod m}} |k(b)| \leq \frac{1}{x} \sum_{0 \leq b < x} |k(b)|.$$

Wegen Satz 12 strebt der letzte Ausdruck gegen 0 bei $x \to \infty$. Daraus folgt die Behauptung. □

Literatur

[1] N.J. Fine: *The distribution of the sum of digits (mod p)*. Bull. AMS **71** (1965), 651–652.

[2] E. Fouvry und C. Mauduit: *Sommes des chiffres et nombres presque premiers*. Math. Ann. **305** (1996), 571–599.

[3] A.O. Gelfond: *Sur les nombres qui ont des propriétés additives et multiplicatives données*. Acta Arithmetica **13** (1968), 259–265.

[4] C. Mauduit und A. Sárközy: *On finite pseudorandom binary sequences II*. J. Number Theory **73** (1998), 256–276.

Jan-Christoph Puchta
Albert-Ludwigs-Universität Freiburg
Mathematisches Institut
Eckerstr. 1
D–79104 Freiburg

Jürgen Spilker
Albert-Ludwigs-Universität Freiburg
Mathematisches Institut
Eckerstr. 1
D–79104 Freiburg
e-mail: spilker@mathematik.uni-freiburg.de