3.4 Die Verteilung der linearen Komplexität

Die Verteilung der linearen Komplexität von Bitfolgen fester Länge lässt sich exakt bestimmen. Bei gegebener Folge $u = (u_0, \ldots, u_{N-1}) \in \mathbb{F}_2^N$ sei $\tilde{u} = (u_0, \ldots, u_N) \in \mathbb{F}_2^{N+1}$ eine Verlängerung um 1 Bit. Die Beziehung zwischen $\lambda(\tilde{u})$ und $\lambda(u)$ wird durch die MASSEY-Rekursion beschrieben: Sei

$$\delta = \begin{cases} 0, & \text{wenn die Vorhersage stimmt,} \\ 1 & \text{sonst;} \end{cases}$$

mit „Vorhersage“ ist der Wert gemeint, den das zu u konstruierte Schieberegister als nächsten liefert. Dann gilt

$$\lambda(\tilde{u}) = \begin{cases} \lambda(u), & \text{wenn } \delta = 0, \\ \lambda(u), & \text{wenn } \delta = 1 \text{ und } \lambda(u) > \frac{N}{2}, \\ N + 1 - \lambda(u), & \text{wenn } \delta = 1 \text{ und } \lambda(u) \leq \frac{N}{2}. \end{cases}$$

Im mittleren Fall wird zwar ein neues Schieberegister benötigt, aber dieses hat dieselbe Länge.

Nun wird eine Formel für die Anzahl aller Folgen der Länge N aufgestellt, die eine gegebene lineare Komplexität l haben. Sei dazu

$$\begin{aligned} M_N(l) &:= \{u \in \mathbb{F}_2^N \mid \lambda(u) = l\} \quad \text{für } N \geq 1 \text{ und } l \in \mathbb{N}, \\ \mu_N(l) &:= \#M_N(l). \end{aligned}$$

Folgende Aussagen sind unmittelbar klar:

- $0 \leq \mu_N(l) \leq 2^N$,
- $\mu_N(l) = 0$ für $l > N$,
- $\sum_{l=0}^{N} \mu_N(l) = 2^N$.

Damit lässt sich nun die Rekursion von $\mu_{N+1}(l)$ auf $\mu_N(l)$ explizit machen.

1. Fall: $0 \leq l \leq \frac{N}{2}$. Jedes $u \in \mathbb{F}_2^N$ hat zwei mögliche Fortsetzungen: $u_N = 0$ oder 1. Genau eine davon stimmt mit der Vorhersage überein und führt zu $\tilde{u} \in M_{N+1}(l)$; die andere führt zu $\tilde{u} \in M_{N+1}(N+1-l)$. Da es keine anderen Beiträge zu $M_{N+1}(l)$ geben kann, folgt $\mu_{N+1}(l) = \mu_N(l)$.

2. Fall: $l = \frac{N+1}{2}$ (was natürlich nur für ungerades N auftreten kann). Das richtig vorhergesagte u_N führt zu $\tilde{u} \in M_{N+1}(l)$, das falsch vorhergesagte wegen der MASSEY-Rekursion aber ebenfalls. Daher folgt $\mu_{N+1}(l) = 2 \cdot \mu_N(l)$.

3. Fall: $l \geq \frac{N}{2} + 1$. Beide möglichen Fortsetzungen führen zu einem $\tilde{u} \in M_{N+1}(l)$. Dazu kommt noch je ein Element von der falsch vorhergesagten Fortsetzung aller $u \in M_{N+1-l}(l)$ aus dem ersten Fall. Also ist $\mu_{N+1}(l) = 2 \cdot \mu_N(l) + \mu_{N+1-l}(l)$.

Zusammengefasst:

Hilfssatz 4 *Die Häufigkeitsfunktion $\mu_N(l)$ für Bitfolgen der Länge N mit linearer Komplexität l erfüllt die Rekursion*

$$\mu_{N+1}(l) = \begin{cases} \mu_N(l), & \text{falls } 0 \le l \le \frac{N}{2}, \\ 2 \cdot \mu_N(l), & \text{falls } l = \frac{N+1}{2}, \\ 2 \cdot \mu_N(l) + \mu_{N+1-l}(l), & \text{falls } l \ge \frac{N}{2} + 1. \end{cases}$$

Daraus lässt sich leicht eine explizite Formel gewinnen:

Satz 2 *Die Häufigkeitsfunktion $\mu_N(l)$ für Bitfolgen der Länge N mit linearer Komplexität l ist gegeben durch*

$$\mu_N(l) = \begin{cases} 1, & \text{falls } l = 0, \\ 2^{2l-1}, & \text{falls } 1 \le l \le \frac{N}{2}, \\ 2^{2(N-l)}, & \text{falls } \frac{N+1}{2} \le l \le N, \\ 0, & \text{falls } l > N. \end{cases}$$

Beweis. Im Fall $n = 1$ ist $M_1(0) = \{(0)\}$, $M_1(1) = \{(1)\}$, also $\mu_1(0) = \mu_1(1) = 1$.

Jetzt wird per Induktion von N auf $N+1$ geschlossen. Der Fall $l = 0$ ist dabei trivial, da $M_{N+1}(0) = \{(0, \ldots, 0)\}$, $\mu_{N+1}(0) = 1$. Nun werden wieder drei Fälle unterschieden:

1. Fall: $1 \le l \le \frac{N}{2}$. Hier ist erst recht $1 \le l \le \frac{N+1}{2}$, und $\mu_{N+1}(l) = \mu_N(l) = 2^{2l-1}$.

2. Fall: $l = \frac{N+1}{2}$ (N ungerade). Hier ist $\mu_N(l) = 2^{2(N-l)}$ und der Exponent $2N - 2l = 2N - N - 1 = N - 1 = 2l - 2$, also $\mu_{N+1}(l) = 2 \cdot 2^{2(N-l)} = 2^{2l-2+1} = 2^{2l-1}$.

3. Fall: $l \ge \frac{N}{2} + 1$. Hier ist wieder $\mu_N(l) = 2^{2(N-l)}$. Für $l' = N + 1 - l$ gilt $l' \le N + 1 - \frac{N}{2} - 1 = \frac{N}{2}$, also $\mu_N(l') = 2^{2l'-1}$. Also folgt $\mu_{N+1}(l) = 2\mu_N(l) + \mu_N(l') = 2^{2N-2l+1} + 2^{2N-2l+1} = 2^{2N-2l+2} = 2^{2(N+1-l)}$.

Damit ist der Beweis vollständig. $\diamond$

Die Tabelle 2 bis $N = 10$ und $l = 10$ ergibt ein interessantes Bild.

Beobachtungen:

- Die Zeile l wird ab $N = 2l$ konstant (rot markiert), die Diagonale jeweils ab $N = 2l - 1$ (blau markiert).

	1	2	3	4	5	6	7	8	9	10	$N \rightarrow$
0	1	1	1	1	1	1	1	1	1	1	
1	1	2	2	2	2	2	2	2	2	2	
2		1	4	8	8	8	8	8	8	8	
3			1	4	16	32	32	32	32	32	
4				1	4	16	64	128	128	128	
5					1	4	16	64	256	512	
6						1	4	16	64	256	
7							1	4	16	64	
8								1	4	16	
9									1	4	
10										1	
l											
$\downarrow$											

Tabelle 2: Die Verteilung der linearen Komplexität

- Jede Spalte N enthält von $l = 1$ bis $l = N$ die Zweierpotenzen 2^k, $k = 0, \ldots, N-1$, jeweils genau einmal, erst die ungeraden Zweierpotenzen (rot) in aufsteigender, dann die geraden (blau) in absteigender Reihenfolge.
- Zu jeder Länge N gibt es jeweils genau eine Folge mit der linearen Komplexität 0 und N.