4.3 Perfekte Pseudozufallsgeneratoren

Es ist jetzt an der Zeit, den Begriff „Zufallsgenerator“ – bzw. genauer gesagt, den Begriff „Pseudozufallsgenerator“) formal zu definieren. Dazu braucht man eine unendliche Parametermenge $M \subseteq \mathbb{N}$; für jeden Parameter $m \in M$ soll eine Instanz des Pseudozufallsgenerators definiert sein. Man denke sich etwa M als eine Menge von BLUM-Zahlen. Es sei $M_n = M \cap [2^{n-1} \ldots 2^n[$ die Menge der n-Bit-Zahlen in M und $I = \{n \in \mathbb{N} \mid M_n \neq \emptyset\}$ der Träger von M. Ferner braucht man ein nichtkonstantes Polynom $g \in \mathbb{N}[X]$.

Ein **Pseudozufallsgenerator mit Parametermenge** M **und Streckungspolynom** g ist eine Familie $G = (G_m)_{m \in M}$ von Funktionen

$$G_m \colon X_m \longrightarrow \mathbb{F}_2^{g(n)} \quad \text{mit } X_m \subseteq \mathbb{F}_2^{k(n)},$$

wobei n die Bitzahl von m ist, so dass es eine (deterministische) polynomiale Schaltnetzfamilie $\tilde{G}$ mit $\tilde{G}_n(m, x) = G_m(x)$ gibt. (Mit anderen Worten: Die Zufallsbits sind effizient berechenbar. Insbesondere ist die Funktion r durch ein Polynom beschränkt.) X_m heißt die Menge der Startwerte zum Parameter m. Jedes G_m streckt also eine $k(n)$-Bit-Folge $x \in X_m$ zu einer $g(n)$-Bit-Folge $G_m(x) \in \mathbb{F}_2^{g(n)}$.

Der BBS-Generator passt als Beispiel so zu dieser Definition: M ist die Menge der BLUM-Zahlen oder eine unendliche Teilmenge davon, $X_m = \mathbb{M}_m$, und $G_m(x) = (b_1(x), \ldots, b_{g(n)}(x))$ mit $b_i(x) = \mathrm{lsb}(x_i)$, wobei $x_0 = x$, $x_i = x_{i-1}^2 \bmod m$.

Ein **polynomialer Test** für den Pseudozufallsgenerator G ist eine (probabilistische) polynomiale Schaltnetzfamilie $C = (C_n)_{n \in \mathbb{N}}$,

$$C_n : \mathbb{F}_2^n \times \mathbb{F}_2^{g(n)} \times \Omega_n \longrightarrow \mathbb{F}_2$$

über einem Wahrscheinlichkeitsraum $\Omega_n \subseteq \mathbb{F}_2^{s(n)}$, wobei $s(n)$ die Anzahl der probabilistischen Eingänge von C_n ist. Die Wahrscheinlichkeit, dass der Test für eine von G erzeugte Folge den Wert 1 errechnet, ist

$$p(G, C, m) = P\{(x, \omega) \in X_m \times \Omega_n \mid C_n(m, G_m(x), \omega) = 1\};$$

die Wahrscheinlichkeit, dass der Test für eine beliebige („echt zufällige“) Folge der gleichen Länge den Wert 1 errechnet, ist

$$\bar{p}(C, m) = P\{(u, \omega) \in \mathbb{F}_2^{g(n)} \times \Omega_n \mid C_n(m, u, \omega) = 1\}.$$

Diese beiden Werte sollten im Idealfall gleich sein. Man sagt, der Pseudozufallsgenerator G **besteht den Test** C, wenn für alle nichtkonstanten Polynome $h \in \mathbb{N}[X]$ die Menge der $m \in M$ mit

$$|p(G, C, m) - \bar{p}(C, m)| \geq \frac{1}{h(n)}$$

dünn in M ist. Der Pseudozufallsgenerator G heißt **perfekt**, wenn er alle polynomialen Tests besteht. D. h., es gibt keinen effizienten Algorithmus, der die von Pseudozufallsgenerator erzeugte Bitfolge von einer „echt zufälligen“ Bitfolge unterscheiden kann.

Zum Nachweis der Perfektheit eines Pseudozufallsgenerators G reicht ein scheinbar schwächerer Test. Sei $G_m(x) = (b_1^{(m)}(x), \ldots, b_{g(n)}^{(m)}(x))$ die von G_m aus dem Startwert x erzeugte Bitfolge. Sei $C = (C_n)_{n\in\mathbb{N}}$ eine polynomiale Schaltnetzfamilie,

$$C_n : \mathbb{F}_2^n \times \mathbb{F}_2^{i_n} \times \Omega_n \longrightarrow \mathbb{F}_2$$

mit $0 \leq i_n \leq g(n)-1$, und sei $h \in \mathbb{N}[X]$ ein nichtkonstantes Polynom. Dann sagt man, C habe einen $\frac{1}{h}$-Vorteil bei der Extrapolation von G, wenn die Menge der Parameter $m \in M$ mit

$$P\{(x,\omega) \in X_m \times \Omega_n \mid C_n(m, b_{j_m+1}^{(m)}(x), \ldots, b_{j_m+i_n}^{(m)}(x), \omega) = b_{j_m}^{(m)}(x)\}$$

$$\geq \frac{1}{2} + \frac{1}{h(n)} \tag{2}$$

für einen Index j_m, $1 \leq j_m \leq g(n) - i_n$ nicht dünn in M ist; das heißt, C kann in genügend vielen Fällen aus einer Teilfolge das vorhergehende Bit mit einem kleinen Vorteil extrapolieren. Man sagt, G besteht den **Extrapolationstest**, wenn es keine solche polynomiale Schaltnetzfamilie gibt, die für irgendein Polynom $h \in \mathbb{N}[X]$ einen $\frac{1}{h}$-Vorteil bei der Extrapolation von G hat.

Zum Beispiel besteht der lineare Kongruenzgenerator den Extrapolationstest nicht.

Hauptsatz 1 [YAOs Kriterium] *Für einen Pseudozufallsgenerator G sind folgende Aussagen äquivalent:*

(i) *G ist perfekt.*

(ii) *G besteht den Extrapolationstest.*

Beweis. „(i) $\Longrightarrow$ (ii)“: Wenn G den Extrapolationstest nicht besteht, gibt es eine polynomiale Schaltnetzfamilie C mit $\frac{1}{h}$-Vorteil bei der Extrapolation von G. Sei $A \subseteq M$ die nicht dünne Menge von Parametern, für die die Ungleichung (2) gilt. Daraus wird ein polynomialer Test $C' = (C'_n)_{n\in\mathbb{N}}$ konstruiert:

$$C'_n(m,u,\omega) = C_n(m, u_{j_m+1}, \ldots, u_{j_m+i_n}, \omega) + u_{j_m} + 1;$$

für $m \in \mathbb{F}_2^n - A$ sei dabei $j_m = 1$ gesetzt (auf diesen Wert kommt es nicht an). Es ist also

$$C'_n(m,u,\omega) = 1 \Longleftrightarrow C_n(m, u_{j_m+1}, \ldots, u_{j_m+i_n}, \omega) = u_{j_m}.$$

Für $m \in A$ folgt

$$p(G, C', m) = P\{C_n(m, b^{(m)}_{j_m+1}(x), \ldots, b^{(m)}_{j_m+i_n}(x), \omega) = b^{(m)}_{j_m}(x)\} \geq \frac{1}{2} + \frac{1}{h(n)}.$$

Dieser Wert ist zu vergleichen mit

$$\bar{p}(C', m) = P\{C_n(m, u_{j_m+1}, \ldots, u_{j_m+i_n}, \omega) = u_{j_m}\}$$

$$= P\{C_n(\ldots) = 0 \text{ und } u_{j_m} = 0\} + P\{C_n(\ldots) = 1 \text{ und } u_{j_m} = 1\}.$$

(Die Summe entspricht einer Zerlegung in zwei disjunkte Teilmengen.) Da hier jeweils die Wahrscheinlichkeit des Zusammentreffens zweier unabhängiger Ereignisse steht, ist

$$\bar{p}(C', m) = \frac{1}{2} P\{C_n(\ldots) = 0\} + \frac{1}{2} P\{C_n(\ldots) = 1\} = \frac{1}{2}.$$

Für $m \in A$ gilt also

$$p(G, C', m) - \bar{p}(C', m) \geq \frac{1}{h(n)}.$$

Daher besteht G den Test C' nicht und ist nicht perfekt.

„(ii) $\Longrightarrow$ (i)“: Sei G nicht perfekt. Dann gibt es einen polynomialen Test C, den G nicht besteht, also ein nichtkonstantes Polynom $h \in \mathbb{N}[X]$ und ein $t \in \mathbb{N}$ mit

$$|p(G, C, m) - \bar{p}(C, m)| \geq \frac{1}{h(n)}$$

für m aus einer nicht dünnen Teilmenge $A \subseteq M$ mit $\#A_n \geq \#M_n / n^t$ für unendlich viele $n \in I$. Für mindestens die Hälfte aller $m \in A_n$ gilt $p(G, C, m) > \bar{p}(C, m)$ oder die umgekehrte Ungleichung; zuerst wird der erste dieser Fälle durchgezogen (bei festem n).

Für $k = 0, \ldots, g(n)$ sei

$$p^k_m = P\{C_n(m, t_1, \ldots, t_k, b^{(m)}_{k+1}(x), \ldots, b^{(m)}_{g(n)}(x), \omega) = 1\},$$

wobei $t_1, \ldots, t_k \in \mathbb{F}_2$ zufällige Bits sind; die Wahrscheinlichkeit wird also in $X_m \times (\mathbb{F}_2^k \times \Omega_n)$ gebildet. Es ist

$$p^0_m = p(G, C, m), \quad p^{g(n)}_m = \bar{p}(C, m),$$

$$\frac{1}{h(n)} \leq p^0_m - p^{g(n)}_m = \sum_{k=1}^{g(n)} (p^{k-1}_m - p^k_m)$$

für die betrachteten $m \in A_n$. Es gibt also ein r_m mit $1 \leq r_m \leq g(n)$, so dass

$$p^{r_m - 1}_m - p^{r_m}_m \geq \frac{1}{g(n) h(n)}.$$

Einer dieser Werte r_m kommt mindestens $(\#M_n/2n^t g(n))$-mal vor; er wird k_n genannt.

Sei $\Omega'_n = \mathbb{F}_2^{k_n} \times \Omega_n$. Die polynomiale Schaltnetzfamilie C', deren deterministische Eingänge aus $A_n \times \mathbb{F}_2^{g(n)-k_n}$ und deren probabilistische Eingänge aus Ω'_n besetzt werden, wird für dieses n so definiert:

$$C'_n(m, u_1, \ldots, u_{g(n)-k_n}, t_1, \ldots, t_{k_n}, \omega) = C_n(m, t, u, \omega) + t_{k_n} + 1.$$

Es ist also

$$C'_n(m, u, t, \omega) = t_{k_n} \Longleftrightarrow C_n(m, t, u, \omega) = 1.$$

Nun ist

$$C'_n(m, b^{(m)}_{k_n+1}(x), \ldots, b^{(m)}_{g(n)}(x), t, \omega) = b^{(m)}_{k_n}(x)$$

$$\Longleftrightarrow \begin{cases} C_n(m, t, b^{(m)}_{k_n+1}(x), \ldots, b^{(m)}_{g(n)}(x), \omega) = 1 \quad \text{und } t_{k_n} = b^{(m)}_{k_n}(x) \\ \text{oder} \\ C_n(m, t, b^{(m)}_{k_n+1}(x), \ldots, b^{(m)}_{g(n)}(x), \omega) = 0 \quad \text{und } t_{k_n} \neq b^{(m)}_{k_n}(x) \end{cases}$$

Beide Möglichkeiten sind jeweils ein Zusammentreffen unabhängiger Ereignisse. Die zweite hat daher die Wahrscheinlichkeit $\frac{1}{2}(1 - p_m^{k_n})$. Die erste ist äquivalent zu

$$C_n(m, t_1, \ldots, t_{k_n-1}, b^{(m)}_{k_n}(x), \ldots, b^{(m)}_{g(n)}(x), \omega) = 1 \text{ und } t_{k_n} = b^{(m)}_{k_n}(x);$$

ihre Wahrscheinlichkeit ist $p_m^{k_n-1}/2$. Zusammen ergibt das

$$P\{C'_n(m, b^{(m)}_{k_n+1}(x), \ldots, b^{(m)}_{g(n)}(x), t, \omega) = b^{(m)}_{k_n}(x)\}$$

$$= \frac{1}{2} + \frac{1}{2}(p_m^{k_n-1} - p_m^{k_n}) \geq \frac{1}{2} + \frac{1}{2g(n)h(n)}$$

für mindestens $\#M_n/2n^t g(n)$ der Parameter $m \in M_n$. Mit $u = t + \mathrm{Grad}(g) + 1$ ist das $\geq \#M_n/n^u$ für unendlich viele $n \in I$.

Im Falle $p(G, C, m) < \bar{p}(C, m)$ für mindestens die Hälfte aller $m \in A_n$ wird analog

$$C'_n(m, u, t, \omega) = C_n(m, t, u, \omega) + t_{k_n}$$

gesetzt; damit klappt der Schluss genauso.

Also besteht G den Extrapolationstest nicht (mit $i_n = g(n) - k_n$ und $j_m = k_n$). $\diamond$

Im Beweis wurde übrigens die Nichtgleichmäßigkeit des Berechnungsmodells verwendet: C'_n hängt von k_n ab, und es wurde kein Algorithmus zur Bestimmung von k_n angegeben.

Der Extrapolationstest wirkt etwas unnatürlich, weil er die erzeugten Bits in umgekehrter Richtung extrapoliert. Das steht im Gegensatz zu den

kryptoanalytischen Verfahren, wo man sich bemüht, Bits *vorher*zusagen. Nun denn:

Sei $C = (C_n)_{n\in\mathbb{N}}$ eine polynomiale Schaltnetzfamilie,

$$C_n : \mathbb{F}_2^n \times \mathbb{F}_2^{i_n} \times \Omega_n \longrightarrow \mathbb{F}_2$$

mit $0 \leq i_n \leq g(n) - 1$, und sei $h \in \mathbb{N}[X]$ ein nichtkonstantes Polynom. Dann hat C einen $\frac{1}{h}$-Vorteil bei der Vorhersage von G, wenn die Menge der Parameter $m \in M$ mit

$$P\{(x,\omega) \mid C_n(m, b_1^{(m)}(x), \ldots, b_{i_n}^{(m)}(x), \omega) = b_{i_n+1}^{(m)}(x)\} \geq \frac{1}{2} + \frac{1}{h(n)}$$

nicht dünn in M ist. Der Pseudozufallsgenerator G besteht den **Vorhersagetest**, wenn keine polynomiale Schaltnetzfamilie einen Vorteil bei der Vorhersage von G hat. Der Beweis von „(i) $\Longrightarrow$ (ii)“ im Hauptsatz 1 lässt sich direkt auf diese Situation adaptieren und ergibt:

Korollar 1 *Jeder perfekte Pseudozufallsgenerator besteht den Vorhersagetest.*

Korollar 2 *Wenn die Quadratrest-Vermutung richtig ist, ist der BBS-Generator perfekt.*

Beweis. Aus Satz 1 ließe sich sonst eine polynomiale Schaltnetzfamilie konstruieren, die die Quadratrest-Eigenschaft für eine nicht dünne Menge von Blum-Zahlen entscheidet. $\diamond$