2.4 Lineare Kongruenzgeneratoren mit unbekanntem Modul

Schwieriger wird es natürlich, wenn der Modul m ebenfalls unbekannt ist und sich auch nicht leicht erraten lässt. Es wird angenommen, dass man nur ein Stück $x_0, x_1, \ldots$ der Folge zur Verfügung hat. Überraschenderweise ist es leichter, zuerst den Multiplikator zu bestimmen. Aus dem folgenden Satz erhält man in wenigen Schritten einen geeigneten Wert dafür, der dann später bei der Suche nach dem Modul hilft. Man erkennt den noetherschen Ansatz in der Form $y_{t+1} \in \mathbb{Z}y_1 + \cdots + \mathbb{Z}y_t$ wieder.

Satz 5 (Plumstead) *Sei (y_i) die Differenzenfolge des linearen Kongruenzgenerators mit erzeugender Funktion $s(x) = ax + b \bmod m$, $m \geq 2$, und Startwert x_0. Sei $y_1 \neq 0$ und t die kleinste Zahl mit $e = \mathrm{ggT}(y_1, \ldots, y_t) \,|\, y_{t+1}$. Dann gilt:*

(i) $t < 1 + {}^2\log m$.

(ii) *Ist $e = c_1y_1 + \cdots + c_ty_t$ mit $c_i \in \mathbb{Z}$ und $a' = (c_1y_2 + \cdots + c_ty_{t+1})/e$, so ($a' \in \mathbb{Z}$ und)*

$$y_{i+1} \equiv a'y_i \pmod{m} \quad \textit{für alle } i.$$

(iii) *Mit $b' = x_1 - a'x_0$ gilt*

$$x_i = a'x_{i-1} + b' \bmod m \quad \textit{für alle } i.$$

Beweis. (i) Ist $e_j = \mathrm{ggT}(y_1, \ldots, y_j)$ kein Teiler von y_{j+1}, so $e_{j+1} \leq e_j/2$. Da $e_1 = |y_1| < m$, folgt $e = e_t < m/2^{t-1}$, also $t - 1 < {}^2\log m$.

(ii) Es ist

$$ae = c_1ay_1 + \cdots + c_tay_t \equiv c_1y_2 + \cdots + c_ty_{t+1} = a'e \pmod{m}.$$

Der größte gemeinsame Teiler d von m und y_1 teilt e nach Hilfssatz 1, also ist auch $d = \mathrm{ggT}(m, e)$. Die Kongruenz wird zuerst durch d geteilt:

$$a\frac{e}{d} \equiv a'\frac{e}{d} \pmod{\bar{m}}$$

mit dem reduzierten Modul $\bar{m} = m/d$. Da e/d zu $\bar{m}$ teilerfremd ist, kann man es wegdividieren:

$$a \equiv a' \pmod{\bar{m}}, \;\; a = a' + k\bar{m}.$$

Also ist $y_{i+1} \equiv ay_i = a'y_i + ky_i\bar{m} \pmod{m}$. Da $d \,|\, y_i$, folgt $y_i \equiv 0$, also $y_{i+1} \equiv a'y_i \pmod{m}$.

(iii) folgt aus Hilfssatz 1 (viii). $\diamond$

Bemerkungen und Beispiele

1. Sei $m = 8397$, $a = 4381$ und $b = 7364$ [Reeds 1977]. Damit wird erzeugt

$$\begin{array}{lll} x_0 = 2134 & & \\ x_1 = 2160 & y_1 = 26 & e_1 = 26 \\ x_2 = 6905 & y_2 = 4745 & e_2 = 13 \\ x_3 = 3778 & y_3 = -3127 & e_3 = 1 \\ x_4 = 8295 & y_4 = 4517 & \end{array}$$

   Es folgt $c_1 = 87542$, $c_2 = -481$, $c_3 = -1$ und $a' = 416881843$.

2. Sei $m = 2^q + 1$, $a = 2^{q-1}$, $b = 2^q$ und $x_0 = 0$. Nach dem Korollar zum folgenden Hilfssatz 2 ist $y_i = (-1)^{i-1} \cdot 2^{q-i+1}$ für $i = 1, \ldots, q+1$ und daher $e_i = 2^{q-i+1}$. Damit ist $t = q+1$. Die Abschätzung für t im Satz 5 ist also scharf, und man braucht tatsächlich $q+3$ der Folgeglieder x_i, also x_0 bis x_{q+2}, um a' zu ermitteln.

Hilfssatz 2 *Die Folge (c_i) in $\mathbb{Z}$ sei durch $c_0 = 0$, $c_i = 2^{i-1} - c_{i-1}$ für $i \geq 1$, definiert. Dann ist*
(i) $c_i = \frac{1}{3} \cdot [2^i - (-1)^i]$ *für alle i,*
(ii) $c_i - 2c_{i-1} = (-1)^{i-1}$ *für alle $i \geq 1$.*

Beweis. (i) zeigt man durch Induktion und (ii) durch direkte Rechnung. ◇

Korollar 1 *Die Folge (x_i) sei von dem linearen Kongruenzgenerator mit Modul $m = 2^q + 1$, Multiplikator $a = 2^{q-1}$, Inkrement $b = 2^q$ und Startwert $x_0 = 0$ erzeugt; (y_i) sei ihre Differenzenfolge. Dann gilt:*
(i) $x_i = c_i \cdot 2^{q-i+1}$ *für $i = 0, \ldots, q+1$,*
(ii) $y_i = (-1)^{i-1} \cdot 2^{q-i+1}$ *für $i = 1, \ldots, q+1$.*

Ein „Ersatzmultiplikator“ a' läßt sich mit Hilfe von Satz 5 also effizient ermitteln. Nun fehlt noch ein Verfahren zur Ermittlung des Moduls m. Dieser wird durch „sukzessive Korrektur“ eingekesselt; im j-ten Schritt wird ein „Ersatzmodul“ m_j und ein „Ersatzmultiplikator“ a_j bestimmt:

- Im ersten Schritt setzt man $m_1 = \infty$ und $a_1 = a'$. [Rechnen mod ∞ soll einfach Rechnen mit ganzen Zahlen bedeuten, und $\mathrm{ggT}(c, \infty)$ soll c sein, wenn $c \neq 0$, und ∞, wenn $c = 0$.]

- Im j-ten Schritt, $j \geq 2$, sei $y'_j := a_{j-1} y_{j-1} \bmod m_{j-1}$. Dann setzt man $m_j = \mathrm{ggT}(m_{j-1}, y'_j - y_j)$ und $a_j = a_{j-1} \bmod m_j$.

Es wird also stets mit den aktuellen Ersatzwerten m_{j-1} und a_{j-1} für m und a eine Voraussage y'_j für y_j gemacht und diese mit dem tatsächlichen Wert y_j verglichen. Stimmen diese beiden Zahlen nicht überein, so unterscheiden sie sich um ein Vielfaches von m; dann werden die Ersatzwerte

korrigiert. Stets gilt $m \mid m_j$. Die j-te Korrektur ändert an der bisherigen Rechnung nichts, denn $y_i \equiv a_j y_{i-1} \pmod{m_j}$ für $i = 2, \ldots, j$, und auch $y_i \equiv a_j y_{i-1} \pmod{m}$ für alle $i \geq 2$. Auch die eigentliche Folge (x_i) erfüllt stets $x_i \equiv a_j x_{i-1} + b_j \pmod{m_j}$ für $i = 1, \ldots, j$ mit $b_j = x_1 - a_j x_0$ nach Hilfssatz 1 (viii).

Im oben gerechneten Beispiel 1 ist

$$\begin{array}{lll} & m_1 = \infty & a_1 = 416881843 \\ y_2' = 10838927918 & m_2 = 10838923173 & a_2 = 416881843 \\ y_3' = 5420327549 & m_3 = 8397 & a_3 = 4381 \end{array}$$

Der Wert m_3 errechnet sich als

$$\mathrm{ggT}(10838923173, 5420330676) = 8397.$$

Da $m_3 \leq 2x_2$, ist $m = m_3$, $a = a_3$ und $b = x_1 - ax_0 \bmod m = 7364$. Wir haben also die korrekten Werte mit zwei Korrekturen gefunden und dabei keine weiteren Folgenglieder gebraucht als die fünf, die schon zur Bestimmung von a' nötig waren. Auffallend sind die großen Zwischenergebnisse, so dass man mit der gewöhnlichen Ganzzahlarithmetik nicht mehr auskommt, sondern eine Arithmetik mit erweiterter Stellenzahl braucht.

Kommt das Verfahren stets zum Ziel? Spätestens wenn die Periode der Folge erreicht ist, also nach höchstens m Schritten, ist die gesamte Folge korrekt voraussagbar. Diese Schranke hat allerdings keinen praktischen Wert. Leider ist sie schon scharf: Bei beliebigem m sei $a = 1$, $b = 1$ und $x_0 = 0$. Dann ist $x_i = i$ und $y_i = 1$ für $i = 0, \ldots, m-1$. Der Startwert für den Ersatzmultiplikator ist $a' = 1$. Die erste falsche Voraussage ist $y_m' = 1$ statt des korrekten Werts $y_m = 1 - m$. Erst nach Auswertung von x_m ist das Verfahren beendet. Nun ist dieser schlechteste Fall leicht erkennbar und separat zu behandeln. Er erschwert aber das Auffinden guter allgemeingültiger Ergebnisse, und in der Tat sind keine solchen bekannt.

Ein etwas anderer Gesichtspunkt ergibt sich, wenn man die Anzahl der notwendigen Korrekturschritte zählt, also die Schritte, in denen der Ersatzmodul sich ändert. Ist nämlich $m_j \neq m_{j-1}$, so $m_j \leq m_{j-1}/2$. Sei $m^{(0)} = \infty > m^{(1)} > \ldots$ die Folge der *verschiedenen* Ersatzmoduln. Dann gilt

$$m^{(1)} = m_{j_1} = |y_{j_1}' - y_{j_1}| < a'|y_{j_1-1}| + m < m(a'+1),$$

$$m \leq m^{(j)} < \frac{m(a'+1)}{2^{j-1}},$$

also stets $j < 1 + {}^2\log(a'+1)$. Damit ist eine obere Schranke für die Anzahl der nötigen Korrekturen gefunden. Joan Plumstead-Boyar gab auch einen Algorithmus an, der zu einem eventuell kleineren Wert von a' und zu der oberen Schranke $2 + {}^2\log m$ für die Anzahl der Korrekturschritte führt. Allerdings wird diese Anzahl von Korrekturschritten in der Regel gar nicht erreicht, so dass die Schranke als Abbruchkriterium nichts nützt.

Hier scheint noch ein lohnendes Betätigungsfeld für die Suche nach theoretischen Ergebnissen offenzustehen. Lässt sich eine kleine Klasse von (vielleicht sowieso schlechten) linearen Kongruenzgeneratoren ausgrenzen, so dass für den großen Rest ein praktisch brauchbares Abbruchkriterium herleitbar ist? Das ist eigentlich zu erwarten. Lässt sich die Verteilung der nötigen Schrittzahl in den Griff bekommen? Wenigstens der Mittelwert? Jedenfalls reichen die vorliegenden Ergebnisse schon, um lineare Kongruenzgeneratoren endgültig als kryptologisch ungeeignet einzustufen.