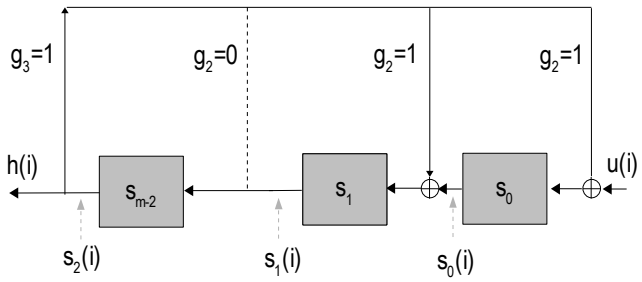
Zur Erzeugung pseudostatistischer Binärfolgen

Für das gemäß des Nachweises von **C. E. Shannnon** perfekt arbeitende symmetrische One-Time-Pad-Verschlüsselungsverfahren von **Vernam** ist die Verwendung binärer Zufallsfolgen erforderlich. Wie bei jedem symmetrischen Verfahren stellt die sichere Schlüsselübertragung aber auch hier eine erhebliche Schwachstelle dar, insbesondere wegen der für jeden Klartext neu zu erzeugenden Zufallsschlüssel gleicher Länge.

Ein wenig, wenn auch in dieser einfachen Form noch nicht entscheidend, lässt sich das Problem verkleinern, wenn man sich die binären Schlüssel mit Pseudo-Zufallsgeneratoren erzeugt. Sie weisen beinahe ideale Autokorrelationsfunktionen auf. Zur Frage, ob damit auch wirklich **statistisch unabhängige** Bitfolgen erzeugt werden, siehe Anhang. Diese Generatoren bestehen im einfachsten Fall aus binären Schieberegistern mit m* Speichern (FlipFlops). Die Länge m* ist dabei der Grad eines definierenden irreduziblen Polynoms g*(x) über Koeffizienten im Z_2, welches als Nullstelle ein **primitives Element** im Galoisfeld GF(2^{m^*}) besitzt. „Primitiv“ ist eine solche Nullstelle dann, wenn sich ihre Potenzen erst ab dem Exponenten $2^{m^*}-1$ wiederholen.

Dieses Generatorpolynom legt mit seinen Koeffizienten die Rückkopplungs-Verbindungen vom Ausgang auf die Speicherelemente fest. Ein Beispiel für das aus der Fehlerkorrektur bekannte irreduzible Polynom

$g^*(x)=x^3+x+1 \rightarrow 1011$ (Nullstelle $\alpha^3=\alpha+1$):



Hält man den rechten Eingang immer auf u(i)=0, belegt wenigstens eines der Speicherelemente mit dem Wert „1“ und startet dann das Schieberegister, so erhält man eine „zufällige“ binäre Folge, die sich erst nach $2^3 – 1 = 7$ Werten wiederholt.

Die Mathematik der Zahlentheorie stellt für beliebig große Polynomgrade passende Polynome bereit, so dass man mit sehr geringem technischen Aufwand Schieberegister zur Erzeugung beinahe „unendlich“ langer Zufallsfolgen herstellen kann.

Zur Übergabe der Schlüssel müssen dann immerhin nicht die Binärfolgen selbst, sondern nur das verwendete Generatorpolynom und der Anfangszustand des Schieberegisters übergeben werden. Die Schwachstelle der Übergabe an sich bleibt leider bestehen (einen idealen Ausweg hieraus bietet aus einem ganz anderen Grund das noch in Entwicklung befindliche Verfahren der Quantenkryptografie).

Noch ein Hinweis: Ein lineares Schieberegister stellt in dieser einfachen Form zunächst nur **eine** Komponente eines Zufallsgenerators dar, da sich bei Kenntnis von $2 \cdot m^*$ hintereinander folgenden Bits die Koeffizienten (Rückkopplungstellen) über ein lineares Gleichungssystem bestimmen lassen, siehe unter anderem bei ***Dankmeier, „Grundkurs Codierung“ 2006, Kapitel 4***. Auch wenn es im Allgemeinen nicht einfach sein wird, die Schlüsselbits zu ermitteln, gibt es doch Angriffs-Verfahren, die hierfür eine von Null verschiedene Wahrscheinlichkeit bieten. Man kann aber durch zusätzliche Maßnahmen diese Aufeinanderfolge verhindern. Weitere Hinweise befinden sich im Anhang.

Aufgabe: „Ermitteln Sie für den Startzustand 100 die Ausgangsfolge des oben skizzierten Schieberegisters. Wie sieht die Folge aus, wenn Sie $g^*(x)=x^3+x^2+1 \rightarrow 1101$ wählen?

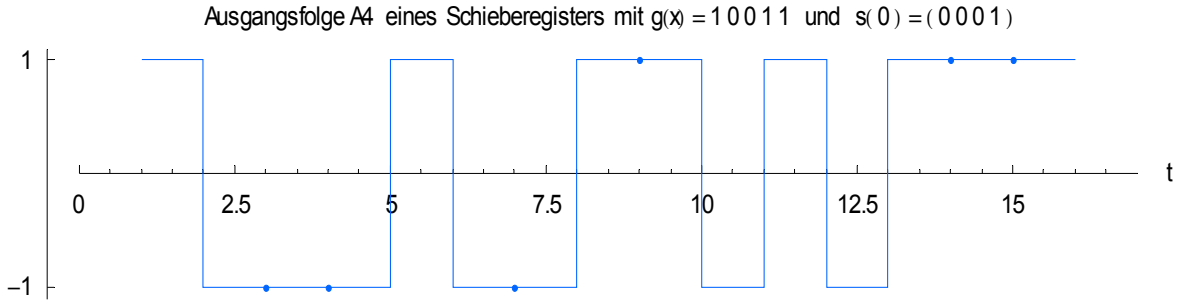
Anhang

Die „Gretchen-Frage“ bleibt, in wie weit die von einem linearen Schieberegister mit der Periode p gelieferte Pseudozufallsfolge wirklich zufällig ist. Dazu muss geklärt werden, was man in diesem Fall unter „zufällig“ versteht. In **„Beutelsbacher/ Neumann/ Schwarzpaul: Kryptografie in Theorie und Praxis, Vieweg, 2010“** (sehr empfehlenswert!) wird in Kapitel 6.2 eine Definition gegeben. Danach erfüllt ein Schieberegister das Kriterium der Nichtvorhersagbarkeit, wenn die Erfolgswahrscheinlichkeit eines Angreifers, das Bit i, i < p, vorherzusagen, höchsten 0.5 beträgt. Dann könnte er auch mit einer Münze werfen.

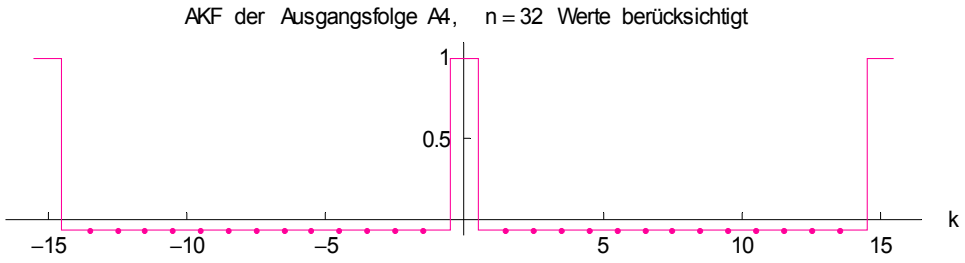
Bisher ist kein Beweis bekannt, der dies für Pseudozufallszahlen-Generatoren feststellen kann. Statt dessen wendet man unterschiedliche statistische Tests an, z. B. den auf Korrelation. Statistisch unabhängige Folgewerte weisen keine Korrelation auf, ihre Autokorrelationsfunktion AKF ist nur im Nullpunkt von 0 verschieden. Für die pseudostatistischen Binärfolgen (auch **PRBS-Folgen** = Pseudo-Random Binary Sequences-Folgen genannt) ist das in sehr guter Näherung erfüllt, wie man es in den folgenden Diagrammen sieht. Der Umkehrschluß, dass dann die Bits statistisch unabhängig wären, gilt jedoch nicht ohne Weiteres. Allgemein: Ist der Nachweis der Zufälligkeit erbracht, so sind auch alle statistischen Tests auf Unabhängigkeit erfüllt. Die Umkehrung gilt leider nicht.

Wenigstens lässt sich zeigen, dass PRBS-Folgen nur schwach korreliert sind. Es besteht also nur ein geringer innerer Zusammenhang zwischen den einzelnen Bits. Das Kriterium der Zufälligkeit ist dennoch nicht erfüllt, da man aus der Kenntnis von 2n aufeinanderfolgenden Bits das Rückkopplungspolynom bestimmen kann.

Betrachten wir ein paar Beispiele. Wird als Rückkopplungspolynom das vom BCH-Code bekannte irreduzible Polynom g(x) = 10011 verwendet, so erhält man bei Zuordnung „0 → +1“ und „1 → -1“ die Folge A4:



A4 hat die AKF (= Autokorrelationsfunktion)



Die AKF wird über $AKF(k) = \frac{1}{2^{m^*}-1} \cdot \sum_{i=1}^{2^{m^*}-1} A4_i \cdot A4_{i+k}, \quad k = 0, 1, 2, \ldots, 2^{m^*}-1$ berechnet.

Jeder Wert der AKF(k) ist das Skalarprodukt der Folge mit ihrer um k verschobenen Version.

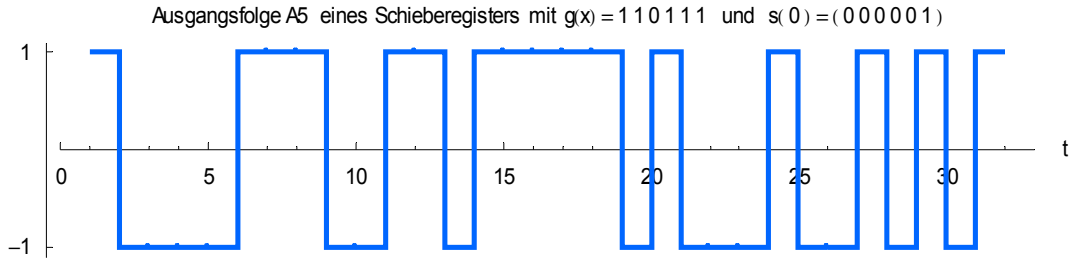
Der Wert AKF(k=0) liefert den Mittelwert der quadratischen Summe (= Quadrat des Effektivwertes, wenn man A4 als Spannungsverlauf interpretiert), die Werte für $k \neq 0$ sind die normalen arithmetischen Mittelwerte und kennzeichnen die Stärke des inneren Zusammenhang der Bits, der hier wegen

$$\frac{1}{2^{(m^*)}-1}$$

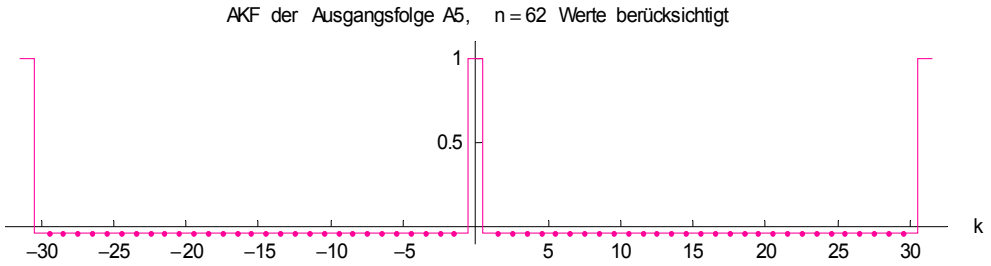
klein bleibt. Binärfolgen mit „+1“ / „-1“ sind praktisch mittelwertfrei. Da wegen der ungeradzahligen Längen die Anzahl der „-1“-Werte um 1 größer als die der „+1“-Werte ist, ergibt sich ein kleiner negativer Rest.

Hinweis: Dies ist z. B. für Radaranwendungen sehr geeignet, siehe Kapitel 4 in „Dankmeier, Grundkurs Codierung, Vieweg 2006“. Man sendet eine PRBS-Impulsfolge und korreliert sie mit den meist stark verrauschten Echos. Die Korrelation - hier genau genommen eine „Kreuz“-Korrelation wegen der zwei unterschiedlichen Signale - lässt den Beitrag der PRBS-Folge unverändert, der Beitrag des Rauschens wird aber reduziert. Als Ergebnis erhält man den Nullpunkt-Peak der PRBS-Folge, der sich über dem „Nebel“ des Rauschanteils erhebt. Aus der Lage des Peaks kann nun die Laufzeit zwischen Senden und Empfangen abgelesen werden. Ohne Korrelation „verschwindet“ die PRBS-Folge im Rauschen. Überhaupt ist Korrelation ein starkes Werkzeug bei der Identifizierung erwarteter Informationen in einem Meer von Daten, z. B. für die Mustererkennung bei Bild- oder Audiodaten.

Zurück zu unserer Anwendung: Mit irreduziblen Polynomen höheren Grades lassen sich längere PRBS-Folgen erzeugen. Z. B. erhält man für ein Schieberegister der Länge 5 diese Sequenz:



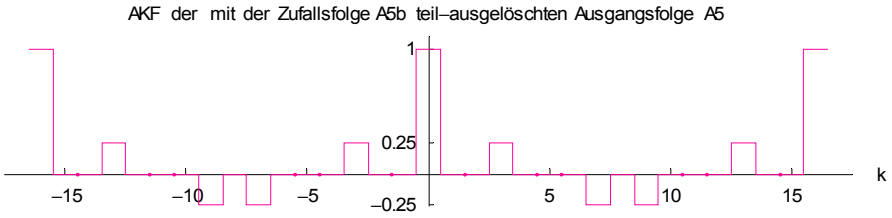
Die AKF ist



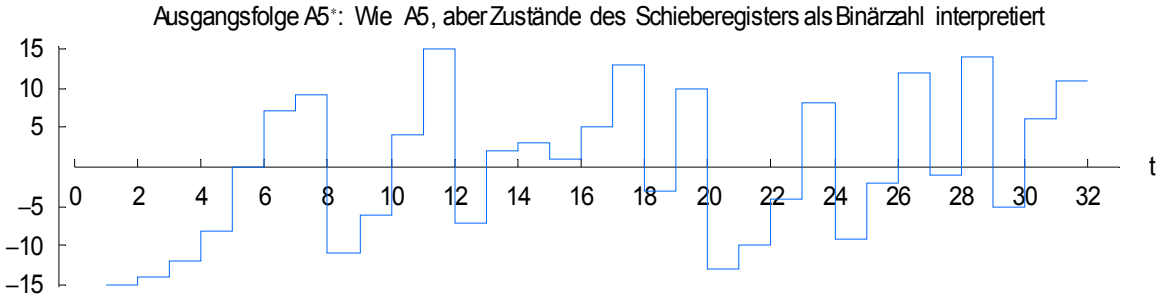
Man ahnt, dass der Verlauf der AKF, von der Folgenlänge abgesehen, gegen den Grad des Rückkopplungspolynoms invariant ist. Z. B. erhält man mit einem irreduziblen Polynom vom Grad 50 eine gleichwertige AKF.

Wenn auch alle Folgen aus derartig aufgebauten linearen Schieberegistern ein hohes Maß an Zusammenhanglosigkeit der Bits garantieren, ist die statistischen Unabhängigkeit trotzdem nicht gegeben, siehe oben. Man kann aber durch geeignete Maßnahmen Verbesserungen erzielen. Z. B. lässt sich die „verräterische“ Aufeinanderfolge der Bits zerstören, in dem man ein zweites Schieberegister B mit einem anderen Rückkopplungspolynom einsetzt und vom ersten Register A nur diejenigen Bits verwendet, bei denen Register B im gleichen Takt eine 1 liefert (auch **Shrinking-Generatoren** benannt).

Im folgenden Beispiel wurde das gemacht und man erhält für die Binärfolge eine AKF, die weiterhin eine nur schwache Korrelation zeigt, aber einem Angreifer nun nicht mehr die Möglichkeit bietet, aufeinanderfolgende Bits zur Bestimmung des Rückkopplungspolynoms zu verwenden. Wenn man das Ergebnis genauer analysiert, konnte zwar immer noch keine perfekte Zufälligkeit erreicht werden, was man auch an den von 0 verschiedenen lokalen Korrelationen sieht, sie wurde aber wesentlich verbessert. Letztlich liefern Operationen, in die ausschließlich periodische Vorgänge eingehen, prinzipiell keine vollkommen statistisch unabhängige Größen.

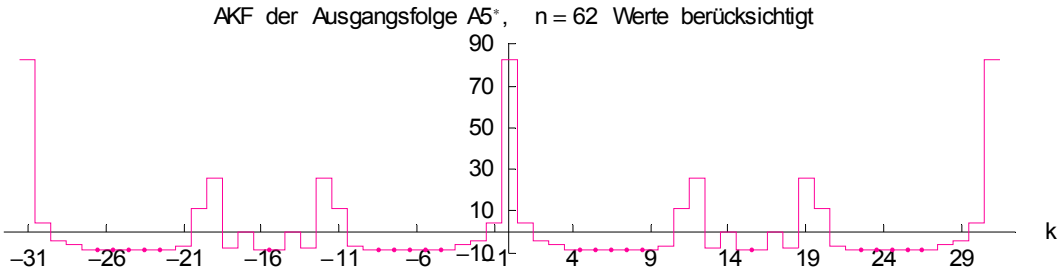


Dass die einfachen Folgen nicht als ausreichend zufällig bezeichnet werden können, sieht man auch, wenn man die Registerzustände - also die Belegung der FlipFlops mit „0“ oder „1“ - als Binärzahl in Dezimaldarstellung interpretiert. Für die zu A5 gehörende mittelwertfreie Folge A5* sieht das so aus:



„Mittelwertfrei“ bedeutet dabei, dass man den Mittelwert der ursprünglichen Folge von dieser abzieht, dabei entstehen aus den ausschließlich positiven Zahlen auch negative Werte, wie das Diagramm zeigt.

Die AKF hat folgenden Verlauf:



Schon die Werte um den Nullpunkt herum fallen nicht sofort auf den kleinen negativen Rest zwischen k= 4 und k=9 ab. In der Mitte zwischen k=10 und k=12, sowie zwischen k=19 und k=21 steigen die Werte sogar an. Die „Zufalls“-Zahlen sind leicht korreliert – und damit nicht statistisch unabhängig.