ALTERNATING STEP GENERATORS CONTROLLED BY DE BRUIJN SEQUENCES

C.G. Günther
Brown Boveri Research Center
5405 Baden, Switzerland

ABSTRACT

The alternating step generator (ASG) is a new generator of pseudorandom sequences which is closely related to the stop-and-go generator. It shares all the good properties of this latter generator without possessing its weaknesses. The ASG consists of three subgenerators K, M, and $\overline{M}$. The main characteristic of its structure is that the output of one of the subgenerators, K, controls the clock of the two others, M and $\overline{M}$. In the present contribution, we determine the period, the distribution of short patterns and a lower bound for the linear complexity of the sequences generated by an ASG. The proof of the lower bound is greatly simplified by assuming that K generates a de Bruijn sequence. Under this and other not very restrictive assumptions the period and the linear complexity are found to be proportional to the period of the de Bruijn sequence. Furthermore the frequency of all short patterns as well as the autocorrelations turn out to be ideal. This means that the sequences generated by the ASG are provably secure against the standard attacks.

1. INTRODUCTION

In stream cipher cryptography messages are usualy combined with pseudorandom sequences by modular addition. Therefore, schemes for the generation of such sequences are important. They are generally based on finite state machines and most frequently on linear feedback shift registers (LFSR's). To avoid certain classes of attacks, these sequences are required to have a large period, a high linear complexity and good statistical properties.

D. Chaum and W.L. Price (Eds.): Advances in Cryptology - EUROCRYPT '87, LNCS 304, pp. 5-14, 1988.

In one approach to the generation of these sequences, the clock of an LFSR is controlled by the output of another LFSR. Examples of generators based on this principle are various kinds of stop-and-go generators [1]-[5] and binary rate multipliers [6]. Both types of generators easily produce sequences of large period and high linear complexity (exponential in the length of the register which controls the clock). The binary rate multipliers furthermore generate sequences with good statistical properties. One disadvantage of these generators is, however, that they need several clock cycles for the generation of one single pseudorandom bit.

Amongst the various kinds of stop-and-go generators we consider the following one:

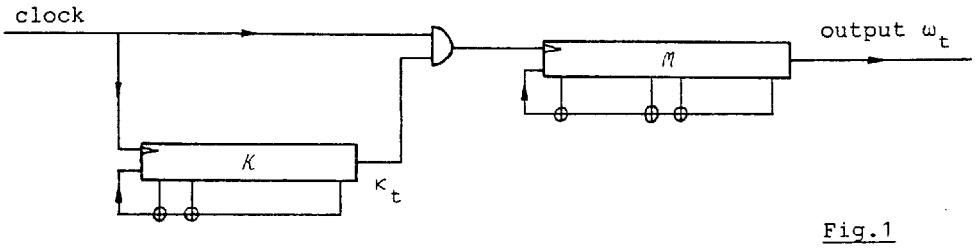

Fig.1

In this generator the output of M is repeated each time the K register produces a "0". On the one side, this leads, under suitable conditions, to a large period and a high linear complexity, on the other side, this always implies bad statistics (eg. $p(00) \cong p(11) \cong \frac{3}{8}$, $p(01) = p(10) \cong \frac{1}{8}$). Furthermore, the fact that the output ω_t can only change if $\kappa_t = 1$, determines one half of all the "1"'s present in the sequence κ. This can strongly reduce the effort needed to reconstruct κ. Similar weaknesses exist in all known stop-and-go generators.

II. THE ALTERNATING STEP GENERATOR

The alternating step generator (ASG) is closely related to the stop-and-go generator. Noteworthy is that it has all the good properties of the latter generator but does not share its weaknesses.

The ASG consists of three subgenerators K, M and $\overline{M}$, which are interconnected such that M and $\overline{M}$ are clocked when the output of K equals "1" and "0", respectively (Fig. 2).

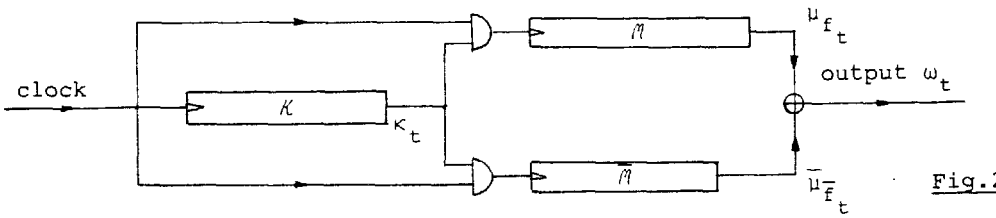

Fig.2

Mathematically, this generator can be described as follows: let κ, μ and $\bar{\mu}$ be the sequences generated by the subgenerators K, M and $\overline{M}$, when they are independently clocked. In addition, let $f_t := \sum_{s=0}^{t-1} \kappa_s$ and $\bar{f}_t = t - f_t$, then the output ω_t is described by

$$\omega_t = \mu_{f_t} \oplus \bar{\mu}_{\bar{f}_t} \quad . \tag{1}$$

In practice the sequences κ, μ amd $\bar{\mu}$ will typically be either maximum length linear recurring sequences (m-sequences) or linear recurring sequences. In the present paper, however, κ will be assumed to be a de Bruijn sequence [7]. Such a sequence can easily be obtained from an m-sequence. In the case of a de Bruijn sequence, the proof for a lower bound on the linear complexity becomes particularly simple. A treatment of the case in which κ, μ and $\bar{\mu}$ are all linear recurring sequences as well as some clues on the cascading of the structure can be found in [8].

The only attack on the ASG we could find so far is a correlation attack on κ [9]. In the present case, however, it does not substan-

tially reduce the effort to break the system. In this correlation attack (Fig. 3) a trial sequence $\tilde{\kappa}$ is correlated with κ using the relation

$$\omega_t \oplus \omega_{t-1} = \begin{cases} \mu_{f_t} \oplus \mu_{f_t-1} \\ \bar{\mu}_{\bar{f}_t} \oplus \bar{\mu}_{\bar{f}_t-1} \end{cases} \quad \text{if} \quad \kappa_{t-1} = \begin{cases} 1 \\ 0 \end{cases} \tag{1}$$

and the fact that the sum of two linear recurring sequences is again such a sequence. The signature for $\tilde{\kappa} = \kappa$ is that the linear complexity of the sequences ν and/or $\bar{\nu}$ (Fig.3) does not increase beyond the length of M and/or $\bar{M}$, respectively. This is determined by the Massey-Berlekamp algorithm.

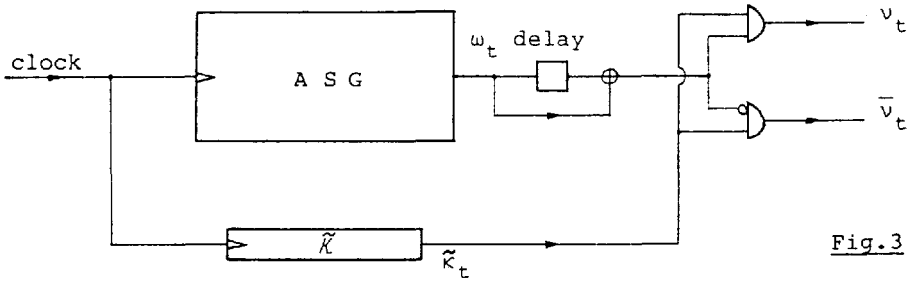

Fig.3

Fortunately, this attack only reduces the effort to break the system to essentially the third root of the effort needed for an exhaustive search. For typical parameters $K = T(\kappa) \sim 2^{127}$ it would need 10^{18} years to search through all phases if 10^{12} phases could be tested per second. In the following section the results on the period, linear complexity and frequencies of short patterns are presented.

III. THE MAIN RESULTS

Theorem 1: (period and linear complexity)

Assume: a) κ is a de Bruijn sequence of period $K = 2^k$,
b) the characteristic polynomials $p(x)$ and $\bar{p}(x)$ of μ and $\bar{\mu}$ are irreducible and different and have the degrees m and $\bar{m}$ and the periods M and $\bar{M}$, respectively,
c) $M, \bar{M} > 1$; $\gcd(M,\bar{M}) = 1$.

Under these assumptions the period T and the linear complexity L of ω satisfy the following relations:

$$T = 2^k M\bar{M} \tag{3}$$

$$(m+\bar{m})2^{k-1} < L \leq (m+\bar{m})2^k \tag{4}$$

Proof: Using that $p(x)$ and $\bar{p}(x)$ are relatively prime, the proof follows immediatly ([10]-[11]) from

i) $S := T(\mu_f) = 2^k M$

ii) the characteristic polynomial of μ_f has the form $p(x)^\ell$ with $2^{k-1} < \ell \leq 2^k$,

and corresponding assertions for $\bar{\mu}_{\bar{f}}$.

The proof of i) only requires $2 \nmid M$, which is implied by the irreducibility of $p(x)$ [10]. It reads as follows: The defining equation of S, i.e. $\mu_{f_{t+S}} = \mu_{f_t}$, $\forall\, t \in Z$, implies $f_{t+S} \equiv f_t \pmod M$, $\forall\, t \in Z$. With $M > 1$ the difference of this equation and of the corresponding equation for $t+1$, i.e. $\kappa_{t+S} \equiv \kappa_t \pmod M$, becomes $\kappa_{t+S} = \kappa_t$, i.e. $S = \gamma 2^k$. As a de Bruijn sequence is "1" with frequency one half, this implies $f_t + \gamma 2^{k-1} \equiv f_t \pmod M$ and as $2 \nmid M$: $\gamma = M$, i.e. $S = 2^k M$.

The proof of ii) is very similar to that of the lower bound for the linear complexity of a de Bruijn sequence [12]. Let D be the time shift operator $D\kappa_t = \kappa_{t-1}$ and let $p(x) = \bigoplus_{i=0}^{m} \pi_i x^i$, then

$$
\begin{aligned}
p(D^{2^k})\mu_{f_t} &= \bigoplus_{i=0}^{m} \pi_i \, \mu_{f_{t-i2^k}} \\
&= \bigoplus_{i=0}^{m} \pi_i \, \mu_{f_t - i2^{k-1}} \qquad (5) \\
&= p(\tilde{D}^{2^{k-1}})\mu_{\tilde{t}} \\
&= p(\tilde{D})^{2^{k-1}} \mu_{\tilde{t}} \\
&= 0 \quad ,
\end{aligned}
$$

where $\tilde{t}$ and $\tilde{D}$ were defined by $\tilde{t} = f_t$ and $\tilde{D}\mu_{\tilde{t}} = \mu_{\tilde{t}-1}$. This equation implies that the characteristic polynomial of μ_{f_t} must divide $p(x^{2^k}) = p(x)^{2^k}$, i.e. it must have the form $p(x)^{\ell}$, with $\ell \leq 2^k$. Now assume $\ell \leq 2^{k-1}$, then

$$
p(x)^{\ell} \mid (x^M-1)^{\ell} \mid (x^M-1)^{2^{k-1}} = x^{2^{k-1}M} - 1 \quad , \qquad (6)
$$

which contradicts $S = 2^k M$. This completes the proof. □

The results of theorem 1 are easily adapted to the case that no assumptions are made on $p(x)$ and $\bar{p}(x)$:

$$
T = 2^k M\bar{M} \qquad (7)
$$

$$
2^{k+1} < L \leq (m+\bar{m})2^k \quad . \qquad (8)
$$

The proof is based on the fact that $\gcd(M,\bar{M}) = 1$ implies $\gcd(p(x),\bar{p}(x)) \mid x-1$ and can easily be figured out.

The following theorem on the frequency of patterns holds for almost arbitrary κ. However, we will restrict ourselves to the case where κ is a de Bruijn sequence, since we would otherwise need a more general assertion on the period. For a more general statement we refer to [8]. In this theorem we use the notation $Z/(T) := \{0,1,\ldots,T-1\}$.

Theorem 2: (frequency of short patterns)

Assume: a) κ is a de Bruijn sequence of period $K = 2^k$,
b) μ and $\bar{\mu}$ are m-sequences with the periods $M = 2^m-1$ and $\bar{M} = 2^{\bar{m}}-1$, respectively,
c) gcd $(M,\bar{M}) = 1$.

Under these assumptions the frequency of any pattern σ of length $\ell \leq \min\{m,\bar{m}\}$ is $2^{-\ell}$ up to an error of order $O(\frac{1}{2^{m-\ell}}) + O(\frac{1}{2^{\bar{m}-\ell}})$, i.e.

$$\frac{1}{T}\,\mathrm{card}\{t\varepsilon Z/(T) \mid \omega_{t+i} = \sigma_i,\ \forall i\varepsilon Z/(\ell)\} = \frac{1}{2^\ell} + O(\frac{1}{2^{m-\ell}}) + O(\frac{1}{2^{\bar{m}-\ell}}) \tag{9}$$

for any $\sigma = (\sigma_0,\ldots,\sigma_{\ell-1}) \varepsilon \{0,1\}^\ell$.

Remark:

We note that the deviation of this distribution from an ideal one is very similar to the corresponding deviation for an m-sequence. In addition, this deviation is due to the corresponding deviation for m-sequences.

Proof of theorem 2: Let $t \varepsilon Z/(T)$ be represented in the form $t = r + (s+\bar{s}M)2^k$, $r \varepsilon Z/(2^k)$, $s \varepsilon Z/(M)$, $\bar{s} \varepsilon Z/(\bar{M})$ and let us first consider the frequency of patterns for a fixed $r \varepsilon Z/(2^k)$. Let $\rho = \rho(r)$ and $\bar{\rho} = \bar{\rho}(r)$ be defined by

$$\rho_0 := 0 \quad , \quad \bar{\rho}_0 := \sigma_0 \quad ,$$

$$\rho_{i+1} := \rho_i \oplus \kappa_{r+i}(\sigma_{i+1} \oplus \sigma_i) \quad , \tag{10}$$

$$\bar{\rho}_{i+1} := \bar{\rho}_i \oplus (1-\kappa_{r+i})(\sigma_{i+1} \oplus \sigma_i) \quad ,$$

for $i \varepsilon Z/(\ell-1)$. Then σ can be decomposed into $(i \varepsilon Z/(\ell))$

$$\sigma_i = \rho_i \oplus \bar{\rho}_i \quad . \tag{11}$$

For the matching condition at time t

$$w_{t+i} = \sigma_i \quad , \quad i \in Z/(\ell) \quad , \tag{12}$$

this implies

$$\mu_{f_{t+i}} \oplus \bar{\mu}_{\bar{f}_{t+i}} = \rho_i \oplus \bar{\rho}_i \quad , \quad i \in Z/(\ell) \quad . \tag{13}$$

Using the following relations

$$\begin{aligned} f_{r+i+1} &= f_{r+i} + \kappa_{r+i} \\ \bar{f}_{r+i+1} &= \bar{f}_{r+i} + 1-\kappa_{r+i} \end{aligned} \qquad i \in Z/(\ell-1) \quad , \tag{14}$$

the sum of equation (13) and of the corresponding equation for i+1 becomes: (i ε Z/(ℓ-1))

$$\begin{aligned} \mu_{f_{t+i+1}} \oplus \mu_{f_{t+i}} &= \rho_{i+1} \oplus \rho_i \\ \bar{\mu}_{\bar{f}_{t+i+1}} \oplus \bar{\mu}_{\bar{f}_{t+i}} &= \bar{\rho}_{i+1} \oplus \bar{\rho}_i \quad . \end{aligned} \tag{15}$$

This has two solutions: (i ε Z/(ℓ))

$$\mu_{f_{t+i}} = \rho_i \quad , \quad \bar{\mu}_{\bar{f}_{t+i}} = \bar{\rho}_i \tag{16a}$$

and

$$\mu_{f_{t+i}} = 1 \oplus \rho_i \quad , \quad \bar{\mu}_{\bar{f}_{t+i}} = 1 \oplus \bar{\rho}_i \quad . \tag{16b}$$

The number of solutions to this equation is equal to the number of occurences of the pattern σ in the sequence $w_{r+(s+\bar{s}M)2^k}$, s ε Z/(M), $\bar{s}$ ε Z/($\bar{M}$), i.e. to the quantity we want to determine.

Without restricting ourselves we consider the solutions of equation (16a). Making use of the fact that κ has the period $K = 2^k$ and that $\sum_{s=0}^{K-1} \kappa_s = 2^{k-1}$, this equation becomes: ($i \in Z/(\ell)$)

$$\mu_{f_{r+i}} + s2^{k-1} = \rho_i \quad , \tag{17}$$

$$\bar{\mu}_{\bar{f}_{r+i}} + (s+\bar{s}M)2^{k-1} = \bar{\rho}_i \quad . \tag{18}$$

Let $\phi_r := f_{r+\ell-1} - f_r$, then the assumptions $2 \nmid M$ and μ an m-sequence imply that equation (17) has $2^{m-\phi_r-1}$ solutions if $\rho \neq 0$. Let $\bar{\phi}_r = \ell-1-\phi_r$, then similarly $2 \nmid \bar{M}$, $\gcd(M,\bar{M}) = 1$ and $\bar{\mu}$ an m-sequence imply that equation (18) has $2^{\bar{m}-\bar{\phi}_r-1}$ solutions if $\bar{\rho} \neq 0$. This remains true for $\rho = 0$ and/or $\bar{\rho} = 0$ if we accept an error of at most $0(\frac{1}{2^{m-\ell}}) + 0(\frac{1}{2^{\bar{m}-\ell}})$. Clearly the same result also holds for equation (16b). Hence the total number of solutions to equation (12) is $2 \cdot 2^{m-\phi_r-1}\, 2^{\bar{m}-\bar{\phi}_r-1} = 2^{m+\bar{m}-\ell}$, which is independent of r. This finally implies that the frequency of the pattern σ is given by

$$\frac{2^{m+\bar{m}-\ell}}{M\bar{M}} + 0(\frac{1}{2^{m-\ell}}) + 0(\frac{1}{2^{\bar{m}-\ell}}) \quad , \tag{19}$$

and thereby yields the assertion. □

IV. CONCLUDING REMARKS

Under suitable assumptions the alternating step generator (ASG) is a simple and very efficient pseudorandom number generator. It is fast and provably satisfies the usual criteria.

The autocorrelations, which were not dealt with in the present paper, are also ideal for a large range of delays ($|\tau| \in Z/(K)$). [8].

The structure of the ASG is favorable to cascading, i.e. to have one or several of the subgenerators K, M and $\bar{M}$ being ASG's themselves. This is further discussed in [8].

SELECTED REFERENCES

[1] S.A. Tretter, "Properties of PN^2 sequences", IEEE Trans. Inform. Theory, vol. IT-20, pp. 295-297, March 1974.

[2] K. Kjeldsen and E. Andresen, "Some randomness properties of cascaded sequences", IEEE Trans. Inform. Theory, vol. IT-26, pp. 227-232, March 1980.

[3] T. Beth and F. Piper, "The stop-and-go-generator", in Proc. of EUROCRYPT 84, Springer Lect. Notes in Comp. Science, vol. 209, pp. 88-92.

[4] R. Vogel, "On the linear complexity of cascaded sequences", in Proc. of EUROCRYPT 84, Springer Lect. Notes in Comp. Science, vol. 209, pp. 99-109.

[5] D. Gollman, "Pseudo random properties of cascade connections of clock controlled shift registers", in Proc. of EUROCRYPT 84, Springer Lect. Notes in Comp. Science, vol. 209, pp. 93-98.

[6] W.G. Chambers and S.M. Jennings, "Linear equivalence of certain BRM shiftregister sequences", Electronics Letters, vol. 20, pp. 1018-1019, Nov. 1984.

[7] N.G. de Bruijn, "A combinatorial problem", Proc. K. Ned. Akad. Wet., vol. 49, pp 758-764, 1946.

[8] C.G. Günther, "Alternating step generators", submitted to IEEE Trans. on Inform. Theory.

[9] T. Siegenthaler, "Correlation-immunity of non-linear combining functions for cryptographic applications", IEEE Trans. on Inform. Theory, vol. IT-30, pp. 776-780, Sept. 1984.

[10] N. Zierler, "Linear recurring sequences", J. Soc. Indust. Appl. Math., vol. 7, pp. 31-48, March 1959.

[11] E.S. Selmer, Linear Recurrence Relations Over Finite Fields, Department of Mathematics, University of Bergen, Norway 1966.

[12] A.H. Chan, R.A. Games and E.L. Key, "On the complexities of de Bruijn sequences", J. of Comb. Theory, Series A, vol. 33, pp. 233-246, 1982.