Helleseth, T. & Kumar, P.V. “Pseudonoise Sequences”
Mobile Communications Handbook
Ed. Suthan S. Suthersan
Boca Raton: CRC Press LLC, 1999

Pseudonoise Sequences

Tor Helleseth
University of Bergen

P. Vijay Kumar
University of Southern California

8.1 Introduction

Pseudonoise sequences (PN sequences), also referred to as pseudorandom sequences, are sequences that are deterministically generated and yet possess some properties that one would expect to find in randomly generated sequences. Applications of PN sequences include signal synchronization, navigation, radar ranging, random number generation, spread-spectrum communications, multipath resolution, cryptography, and signal identification in multiple-access communication systems. The *correlation* between two sequences $\{x(t)\}$ and $\{y(t)\}$ is the complex inner product of the first sequence with a shifted version of the second sequence. The correlation is called 1) an autocorrelation if the two sequences are the same, 2) a crosscorrelation if they are distinct, 3) a periodic correlation if the shift is a cyclic shift, 4) an aperiodic correlation if the shift is not cyclic, and 5) a partial-period correlation if the inner product involves only a partial segment of the two sequences. More precise definitions are given subsequently.

Binary ***m* sequences**, defined in the next section, are perhaps the best-known family of PN sequences. The balance, run-distribution, and autocorrelation properties of these sequences mimic those of random sequences. It is perhaps the random-like correlation properties of PN sequences that makes them most attractive in a communications system, and it is common to refer to any collection of low-correlation sequences as a family of PN sequences.

Section 8.2 begins by discussing m sequences. Thereafter, the discussion continues with a description of sequences satisfying various correlation constraints along the lines of the accompanying

self-explanatory figure, Fig. 8.1. Expanded tutorial discussions on pseudorandom sequences may be found in [14], in [15, Chapter 5] and in [6].

8.2 *m* Sequences

A binary {0, 1} **shift-register sequence** $\{s(t)\}$ is a sequence that satisfies a linear recurrence relation of the form

$$\sum_{i=0}^{r} f_i s(t+i) = 0\,, \qquad \text{for all } t \geq 0 \tag{8.1}$$

where $r \geq 1$ is the *degree* of the recursion; the coefficients f_i belong to the finite field $GF(2) = \{0, 1\}$ where the leading coefficient $f_r = 1$. Thus, both sequences $\{a(t)\}$ and $\{b(t)\}$ appearing in Fig. 8.2 are shift-register sequences. A sequence satisfying a recursion of the form in Eq. (8.1) is said to have *characteristic polynomial* $f(x) = \sum_{i=0}^{r} f_i x^i$. Thus, $\{a(t)\}$ and $\{b(t)\}$ have characteristic polynomials given by $f(x) = x^3 + x + 1$ and $f(x) = x^3 + x^2 + 1$, respectively.

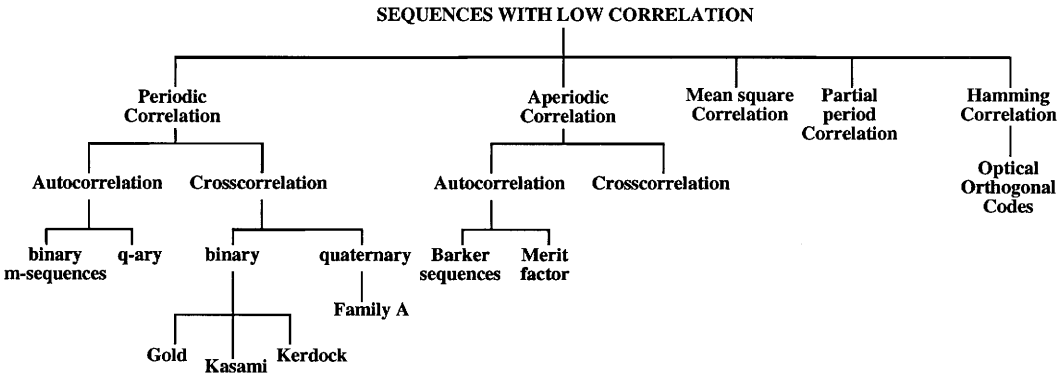

FIGURE 8.1: Overview of pseudonoise sequences.

Since an r-bit binary shift register can assume a maximum of 2^r different states, it follows that every shift-register sequence $\{s(t)\}$ is eventually periodic with period $n \leq 2^r$, i.e.,

$$s(t) = s(t+n), \qquad \text{for all } t \geq N$$

for some integer N. In fact, the maximum period of a shift-register sequence is $2^r - 1$, since a shift register that enters the all-zero state will remain forever in that state. The upper shift register in Fig. 8.2 when initialized with starting state 0 0 1 generates the periodic sequence $\{a(t)\}$ given by

$$0010111 \quad 0010111 \quad 0010111 \quad \cdots \tag{8.2}$$

of period $n = 7$. It follows then that this shift register generates sequences of maximal period starting from any nonzero initial state.

An m sequence is simply a binary shift-register sequence having maximal period. For every $r \geq 1$, m sequences are known to exist. The periodic **autocorrelation** function θ_s of a binary {0, 1} sequence

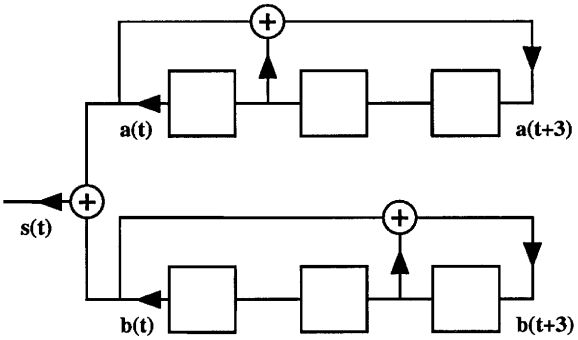

FIGURE 8.2: An example Gold sequence generator. Here $\{a(t)\}$ and $\{b(t)\}$ are m sequences of length 7.

$\{s(t)\}$ of period n is defined by

$$\theta_s(\tau) = \sum_{t=0}^{n-1} (-1)^{s(t+\tau)-s(t)}, \qquad 0 \le \tau \le n-1$$

An m sequence of length $2^r - 1$ has the following attributes. 1) *Balance property:* in each period of the m sequence there are 2^{r-1} ones and $2^{r-1} - 1$ zeros. 2) *Run property:* every nonzero binary s-tuple, $s \le r$ occurs 2^{r-s} times, the all-zero s-tuple occurs $2^{r-s} - 1$ times. 3) *Two-level autocorrelation function:*

$$\theta_s(\tau) = \begin{cases} n & \text{if } \tau = 0 \\ -1 & \text{if } \tau \neq 0 \end{cases} \tag{8.3}$$

The first two properties follow immediately from the observation that every nonzero r-tuple occurs precisely once in each period of the m sequence. For the third property, consider the difference sequence $\{s(t+\tau) - s(t)\}$ for $\tau \neq 0$. This sequence satisfies the same recursion as the m sequence $\{s(t)\}$ and is clearly not the all-zero sequence. It follows, therefore, that $\{s(t+\tau) - s(t)\} \equiv \{s(t+\tau')\}$ for some τ', $0 \le \tau' \le n-1$, i.e., is a different cyclic shift of the m sequence $\{s(t)\}$. The balance property of the sequence $\{s(t+\tau')\}$ then gives us attribute 3. The m sequence $\{a(t)\}$ in Eq. (8.2) can be seen to have the three listed properties.

If $\{s(t)\}$ is any sequence of period n and d is an integer, $1 \le d \le n$, then the mapping $\{s(t)\} \to \{s(dt)\}$ is referred to as a *decimation* of $\{s(t)\}$ by the integer d. If $\{s(t)\}$ is an m sequence of period $n = 2^r - 1$ and d is an integer relatively prime to $2^r - 1$, then the decimated sequence $\{s(dt)\}$ clearly also has period n. Interestingly, it turns out that the sequence $\{s(dt)\}$ is always also an m sequence of the same period. For example, when $\{a(t)\}$ is the sequence in Eq. (8.2), then

$$a(3t) = 0011101 \quad 0011101 \quad 0011101 \quad \cdots \tag{8.4}$$

and

$$a(2t) = 0111001 \quad 0111001 \quad 0111001 \quad \cdots \tag{8.5}$$

The sequence $\{a(3t)\}$ is also an m sequence of period 7, since it satisfies the recursion

$$s(t+3) + s(t+2) + s(t) = 0 \qquad \text{for all } t$$

of degree $r = 3$. In fact $\{a(3t)\}$ is precisely the sequence labeled $\{b(t)\}$ in Fig. 8.2. The sequence $\{a(2t)\}$ is simply a cyclically shifted version of $\{a(t)\}$ itself; this property holds in general. If $\{s(t)\}$ is any m sequence of period $2^r - 1$, then $\{s(2t)\}$ will always be a shifted version of the same m sequence. Clearly, the same is true for decimations by any power of 2.

Starting from an m sequence of period $2^r - 1$, it turns out that one can generate all m sequences of the same period through decimations by integers d relatively prime to $2^r - 1$. The set of integers d, $1 \leq d \leq 2^r - 1$ satisfying $(d, 2^r - 1) = 1$ forms a group under multiplication modulo $2^r - 1$, with the powers $\{2^i \mid 0 \leq i \leq r - 1\}$ of 2 forming a subgroup of order r. Since decimation by a power of 2 yields a shifted version of the same m sequence, it follows that the number of distinct m sequences of period $2^r - 1$ is $[\phi(2^r - 1)/r]$ where $\phi(n)$ denotes the number of integers d, $1 \leq d \leq n$, relatively prime to n. For example, when $r = 3$, there are just two cyclically distinct m sequences of period 7, and these are precisely the sequences $\{a(t)\}$ and $\{b(t)\}$ discussed in the preceding paragraph. Tables provided in [12] can be used to determine the characteristic polynomial of the various m sequences obtainable through the decimation of a single given m sequence. The classical reference on m sequences is [4].

If one obtains a sequence of some large length n by repeatedly tossing an unbiased coin, then such a sequence will very likely satisfy the balance, run, and autocorrelation properties of an m sequence of comparable length. For this reason, it is customary to regard the extent to which a given sequence possesses these properties as a measure of randomness of the sequence. Quite apart from this, in many applications such as signal synchronization and radar ranging, it is desirable to have sequences $\{s(t)\}$ with low autocorrelation sidelobes i.e., $|\theta_s(\tau)|$ is small for $\tau \neq 0$. Whereas m sequences are a prime example, there exist other methods of constructing binary sequences with low out-of-phase autocorrelation.

Sequences $\{s(t)\}$ of period n having an autocorrelation function identical to that of an m sequence, i.e., having θ_s satisfying Eq. (8.3) correspond to well-studied combinatorial objects known as *cyclic Hadamard difference sets.* Known infinite families fall into three classes 1) Singer and Gordon, Mills and Welch, 2) quadratic residue, and 3) twin-prime difference sets. These correspond, respectively, to sequences of period n of the form $n = 2^r - 1$, $r \geq 1$; n prime; and $n = p(p + 2)$ with both p and $p + 2$ being prime in the last case. For a detailed treatment of cyclic difference sets, see [2]. A recent observation by Maschietti in [9] provides additional families of cyclic Hadamard difference sets that also correspond to sequences of period $n = 2^r - 1$.

8.3 The q-ary Sequences with Low Autocorrelation

As defined earlier, the autocorrelation of a binary $\{0, 1\}$ sequence $\{s(t)\}$ leads to the computation of the inner product of an $\{-1, +1\}$ sequence $\{(-1)^{s(t)}\}$ with a cyclically shifted version $\{(-1)^{s(t+\tau)}\}$ of itself. The $\{-1, +1\}$ sequence is transmitted as a phase shift by either 0° and 180° of a radio-frequency carrier, i.e., using binary phase-shift keying (PSK) modulation. If the modulation is q-ary PSK, then one is led to consider sequences $\{s(t)\}$ with symbols in the set Z_q, i.e., the set of integers modulo q. The relevant autocorrelation function $\theta_s(\tau)$ is now defined by

$$\theta_s(\tau) = \sum_{t=0}^{n-1} \omega^{s(t+\tau)-s(t)}$$

where n is the period of $\{s(t)\}$ and ω is a complex primitive qth root of unity. It is possible to construct sequences $\{s(t)\}$ over Z_q whose autocorrelation function satisfies

$$\theta_s(\tau) = \begin{cases} n & \text{if } \tau = 0 \\ 0 & \text{if } \tau \neq 0 \end{cases}$$

For obvious reasons, such sequences are said to have an *ideal autocorrelation function*.

We provide without proof two sample constructions. The sequences in the first construction are given by

$$s(t) = \begin{cases} t^2/2 \pmod{n} & \text{when } n \text{ is even} \\ t(t+1)/2 \pmod{n} & \text{when } n \text{ is odd} \end{cases}$$

Thus, this construction provides sequences with ideal autocorrelation for any period n. Note that the size q of the sequence symbol alphabet equals n when n is odd and $2n$ when n is even.

The second construction also provides sequences over Z_q of period n but requires that n be a perfect square. Let $n = r^2$ and let π be an arbitrary permutation of the elements in the subset $\{0, 1, 2, \ldots, (r-1)\}$ of Z_n: Let g be an arbitrary function defined on the subset $\{0, 1, 2, \ldots, r-1\}$ of Z_n. Then any sequence of the form

$$s(t) = rt_1\pi(t_2) + g(t_2) \pmod{n}$$

where $t = rt_1 + t_2$ with $0 \le t_1, t_2 \le r-1$ is the base-r decomposition of t, has an ideal autocorrelation function. When the alphabet size q equals or divides the period n of the sequence, ideal-autocorrelation sequences also go by the name *generalized bent functions*. For details, see [6].

8.4 Families of Sequences with Low Crosscorrelation

Given two sequences $\{s_1(t)\}$ and $\{s_2(t)\}$ over Z_q of period n, their **crosscorrelation** function $\theta_{1,2}(\tau)$ is defined by

$$\theta_{1,2}(\tau) = \sum_{t=0}^{n-1} \omega^{s_1(t+\tau)-s_2(t)}$$

where ω is a primitive qth root of unity. The crosscorrelation function is important in code-division multiple-access (CDMA) communication systems. Here, each user is assigned a distinct signature sequence and to minimize interference due to the other users, it is desirable that the signature sequences have pairwise, low values of crosscorrelation function. To provide the system in addition with a self-synchronizing capability, it is desirable that the signature sequences have low values of the autocorrelation function as well.

Let $\mathcal{F} = \{\{s_i(t)\} \mid 1 \le i \le M\}$ be a family of M sequences $\{s_i(t)\}$ over Z_q each of period n. Let $\theta_{i,j}(\tau)$ denote the crosscorrelation between the ith and jth sequence at shift τ, i.e.,

$$\theta_{i,j}(\tau) = \sum_{t=0}^{n-1} \omega^{s_i(t+\tau)-s_j(t)}, \qquad 0 \le \tau \le n-1$$

The classical goal in sequence design for CDMA systems has been minimization of the parameter

$$\theta_{\max} = \max\left\{\left|\theta_{i,j}(\tau)\right| \mid \text{ either } i \ne j \text{ or } \tau \ne 0\right\}$$

for fixed n and M. It should be noted though that, in practice, because of data modulation the correlations that one runs into are typically of an aperiodic rather than a periodic nature (see Section 8.5). The problem of designing for low aperiodic correlation, however, is a more difficult one. A typical approach, therefore, has been to design based on periodic correlation, and then to analyze the resulting design for its aperiodic correlation properties. Again, in many practical systems, the mean square correlation properties are of greater interest than the worst-case correlation represented by a parameter such as $\theta_{\max}$. The mean square correlation is discussed in Section 8.6.

Bounds on the minimum possible value of $\theta_{\max}$ for given period n, family size M, and alphabet size q are available that can be used to judge the merits of a particular sequence design. The most efficient bounds are those due to Welch, Sidelnikov, and Levenshtein, see [6]. In CDMA systems, there is greatest interest in designs in which the parameter $\theta_{\max}$ is in the range $\sqrt{n} \leq \theta_{\max} \leq 2\sqrt{n}$. Accordingly, Table 8.1 uses the Welch, Sidelnikov, and Levenshtein bounds to provide an order-of-magnitude upper bound on the family size M for certain $\theta_{\max}$ in the cited range.

Practical considerations dictate that q be small. The bit-oriented nature of electronic hardware makes it preferable to have q a power of 2. With this in mind, a description of some efficient sequence families having low auto- and crosscorrelation values and alphabet sizes $q = 2$ and $q = 4$ are described next.

TABLE 8.1 Bounds on Family Size M for Given n, $\theta_{\max}$

$\theta_{\max}$	Upper bound on M $q = 2$	Upper Bound on M $q > 2$
$\sqrt{n}$	$n/2$	n
$\sqrt{2n}$	n	$n^2/2$
$2\sqrt{n}$	$3n^2/10$	$n^3/2$

8.4.1 Gold and Kasami Sequences

Given the low autocorrelation sidelobes of an m sequence, it is natural to attempt to construct families of low correlation sequences starting from m sequences. Two of the better known constructions of this type are the families of Gold and Kasami sequences.

Let r be odd and $d = 2^k + 1$ where k, $1 \leq k \leq r - 1$, is an integer satisfying $(k, r) = 1$. Let $\{s(t)\}$ be a cyclic shift of an m sequence of period $n = 2^r - 1$ that satisfies $S(dt) \not\equiv 0$ and let $\mathcal{G}$ be the *Gold* family of $2^r + 1$ sequences given by

$$\mathcal{G} = \{s(t)\} \cup \{s(dt)\} \cup \{\{s(t) + s(d[t + \tau])\} \mid 0 \leq \tau \leq n - 1\}$$

Then each sequence in $\mathcal{G}$ has period $2^r - 1$ and the maximum-correlation parameter $\theta_{\max}$ of $\mathcal{G}$ satisfies

$$\theta_{\max} \leq \sqrt{2^{r+1}} + 1$$

An application of the Sidelnikov bound coupled with the information that $\theta_{\max}$ must be an odd integer yields that for the family $\mathcal{G}$, $\theta_{\max}$ is as small as it can possibly be. In this sense the family $\mathcal{G}$ is an optimal family. We remark that these comments remain true even when d is replaced by the integer $d = 2^{2k} - 2^k + 1$ with the conditions on k remaining unchanged.

The Gold family remains the best-known family of m sequences having low crosscorrelation. Applications include the Navstar Global Positioning System whose signals are based on Gold sequences.

The family of Kasami sequences has a similar description. Let $r = 2v$ and $d = 2^v + 1$. Let $\{s(t)\}$ be a cyclic shift of an m sequence of period $n = 2^r - 1$ that satisfies $s(dt) \not\equiv 0$, and consider the family of Kasami sequences given by

$$\mathcal{K} = \{s(t)\} \cup \left\{\{s(t) + s(d[t + \tau])\} \mid 0 \leq \tau \leq 2^v - 2\right\}$$

Then the Kasami family $\mathcal{K}$ contains 2^v sequences of period $2^r - 1$. It can be shown that in this case

$$\theta_{\max} = 1 + 2^v$$

This time an application of the Welch bound and the fact that $\theta_{\max}$ is an integer shows that the Kasami family is optimal in terms of having the smallest possible value of $\theta_{\max}$ for given n and M.

8.4.2 Quaternary Sequences with Low Crosscorrelation

The entries in Table 8.1 suggest that nonbinary (i.e., $q > 2$) designs may be used for improved performance. A family of quaternary sequences that outperform the Gold and Kasami sequences is now discussed below.

Let $f(x)$ be the characteristic polynomial of a binary m sequence of length $2^r - 1$ for some integer r. The coefficients of $f(x)$ are either 0 or 1. Now, regard $f(x)$ as a polynomial over Z_4 and form the product $(-1)^r f(x) f(-x)$. This can be seen to be a polynomial in x^2. Define the polynomial $g(x)$ of degree r by setting $g(x^2) = (-1)^r f(x) f(-x)$. Let $g(x) = \sum_{i=0}^{r} g_i x^i$ and consider the set of all quaternary sequences $\{a(t)\}$ satisfying the recursion $\sum_{i=0}^{r} g_i a(t+i) = 0$ for all t.

It turns out that with the exception of the all-zero sequence, all of the sequences generated in this way have period $2^r - 1$. Thus, the recursion generates a family $\mathcal{A}$ of $2^r + 1$ cyclically distinct quaternary sequences. Closer study reveals that the maximum correlation parameter $\theta_{\max}$ of this family satisfies $\theta_{\max} \leq 1 + \sqrt{2^r}$. Thus, in comparison to the family of Gold sequences, the family $\mathcal{A}$ offers a lower value of $\theta_{\max}$ (by a factor of $\sqrt{2}$) for the same family size. In comparison to the set of Kasami sequences, it offers a much larger family size for the same bound on $\theta_{\max}$. Family $\mathcal{A}$ sequences may be found discussed in [16, 3].

We illustrate with an example. Let $f(x) = x^3 + x + 1$ be the characteristic polynomial of the m sequence $\{a(t)\}$ in Eq. (8.1). Then over Z_4

$$g\left(x^2\right) = (-1)^3 f(x) f(-x) = x^6 + 2x^4 + x^2 + 3$$

so that $g(x) = x^3 + 2x^2 + x + 3$. Thus, the sequences in family $\mathcal{A}$ are generated by the recursion $s(t+3) + 2s(t+2) + s(t+1) + 3s(t) = 0 \bmod 4$. The corresponding shift register is shown in Fig. 8.3. By varying initial conditions, this shift register can be made to generate nine cyclically distinct sequences, each of length 7. In this case $\theta_{\max} \leq 1 + \sqrt{8}$.

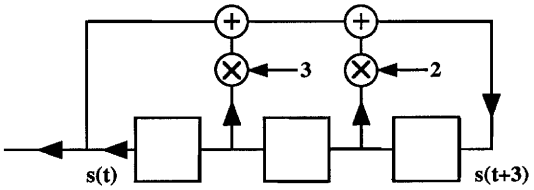

FIGURE 8.3: Shift register that generates family A quaternary sequences $\{s(t)\}$ of period 7.

8.4.3 Binary Kerdock Sequences

The Gold and Kasami families of sequences are closely related to binary linear cyclic codes. It is well known in coding theory that there exists nonlinear binary codes whose performance exceeds that of the best possible linear code. Surprisingly, some of these examples come from binary codes, which are images of *linear quaternary* ($q = 4$) *codes* under the Gray map: $0 \to 00$, $1 \to 01$, $2 \to 11$,

$3 \rightarrow 10$. A prime example of this is the Kerdock code, which recently has been shown to be the Gray image of a quaternary linear code. Thus, it is not surprising that the Kerdock code yields binary sequences that significantly outperform the family of Kasami sequences.

The Kerdock sequences may be constructed as follows: let $f(x)$ be the characteristic polynomial of an m sequence of period $2^r - 1, r$ odd. As before, regarding $f(x)$ as a polynomial over Z_4 (which happens to have $\{0, 1\}$ coefficients), let the polynomial $g(x)$ over Z_4 be defined via $g(x^2) = -f(x)f(-x)$. [Thus, $g(x)$ is the characteristic polynomial of a family $\mathcal{A}$ sequence set of period $2^r - 1$.] Set $h(x) = -g(-x) = \sum_{i=0}^{r} h_i x^i$, and let S be the set of all Z_4 sequences satisfying the recursion $\sum_{i=0}^{r} h_i s(t+i) = 0$. Then S contain 4^r-distinct sequences corresponding to all possible distinct initializations of the shift register.

Let T denote the subset S of size 2^r-consisting of those sequences corresponding to initializations of the shift register only using the symbols 0 and 2 in Z_4. Then the set $S - T$ of size $4^r - 2^r$ contains a set $\mathcal{U}$ of 2^{r-1} cyclically distinct sequences each of period $2(2^r - 1)$. Given $x = a + 2b \in Z_4$ with $a, b \in \{0, 1\}$, let μ denote the most significant bit (MSB) map $\mu(x) = b$. Let $\mathcal{K}_E$ denote the family of 2^{r-1} binary sequences obtained by applying the map μ to each sequence in $\mathcal{U}$. It turns out that each sequence in $\mathcal{U}$ also has period $2(2^r - 1)$ and that, furthermore, for the family $\mathcal{K}_E$, $\theta_{\max} \leq 2 + \sqrt{2^{r+1}}$. Thus, $\mathcal{K}_E$ is a much larger family than the Kasami family, while having almost exactly the same value of $\theta_{\max}$.

For example, taking $r = 3$ and $f(x) = x^3 + x + 1$, we have from the previous family $\mathcal{A}$ example that $g(x) = x^3 + 2x^2 + x + 3$, so that $h(x) = -g(-x) = x^3 + 2x^2 + x + 1$. Applying the MSB map to the head of the shift register, and discarding initializations of the shift register involving only 0's and 2's yields a family of four cyclically distinct binary sequences of period 14. Kerdock sequences are discussed in [6, 11, 1, 17].

8.5 Aperiodic Correlation

Let $\{x(t)\}$ and $\{y(t)\}$ be complex-valued sequences of length (or period) n, not necessarily distinct. Their *aperiodic correlation* values $\{\rho_{x,y}(\tau)| \; -(n-1) \leq \tau \leq n-1\}$ are given by

$$\rho_{x,y}(\tau) = \sum_{t=\max\{0,-\tau\}}^{\min\{n-1,n-1-\tau\}} x(t+\tau)y^*(t)$$

where $y^*(t)$ denotes the complex conjugate of $y(t)$. When $x \equiv y$, we will abbreviate and write ρ_x in place of $\rho_{x,y}$. The sequences described next are perhaps the most famous example of sequences with low-aperiodic autocorrelation values.

8.5.1 Barker Sequences

A binary $\{-1, +1\}$ sequence $\{s(t)\}$ of length n is said to be a *Barker sequence* if the aperiodic autocorrelation values $\rho_s(\tau)$ satisfy $|\rho_s(\tau)| \leq 1$ for all τ, $-(n-1) \leq \tau \leq n-1$. The Barker property is preserved under the following transformations:

$$s(t) \rightarrow -s(t), \qquad s(t) \rightarrow (-1)^t s(t) \quad \text{and} \quad s(t) \rightarrow s(n-1-t)$$

as well as under compositions of the preceding transformations. Only the following Barker sequences are known:

$$\begin{aligned} n &= 2 \quad ++ \\ n &= 3 \quad ++- \\ n &= 4 \quad +++- \\ n &= 5 \quad +++-+ \\ n &= 7 \quad +++--+- \\ n &= 11 \quad +++---+--+- \\ n &= 13 \quad +++++--++-+-+ \end{aligned}$$

where $+$ denotes $+1$ and $-$ denotes -1 and sequences are generated from these via the transformations already discussed. It is known that if any other Barker sequence exists, it must have length $n > 1{,}898{,}884$, that is a multiple of 4.

For an upper bound to the maximum out-of-phase aperiodic autocorrelation of an m sequence, see [13].

8.5.2 Sequences with High Merit Factor

The *merit factor* F of a $\{-1, +1\}$ sequence $\{s(t)\}$ is defined by

$$F = \frac{n^2}{2\sum_{\tau=1}^{n-1} \rho_s^2(\tau)}$$

Since $\rho_s(\tau) = \rho_s(-\tau)$ for $1 \le |\tau| \le n-1$ and $\rho_s(0) = n$, factor F may be regarded as the ratio of the square of the in-phase autocorrelation, to the sum of the squares of the out-of-phase aperiodic autocorrelation values. Thus, the merit factor is one measure of the aperiodic autocorrelation properties of a binary $\{-1, +1\}$ sequence. It is also closely connected with the signal to self-generated noise ratio of a communication system in which coded pulses are transmitted and received.

Let F_n denote the largest merit factor of any binary $\{-1, +1\}$ sequence of length n. For example, at length $n = 13$, the Barker sequence of length 13 has a merit factor $F = F_{13} = 14.08$. Assuming a certain ergodicity postulate it was established by Golay that $\lim_{n\to\infty} F_n = 12.32$. Exhaustive computer searches carried out for $n \le 40$ have revealed the following.

1. For $1 \le n \le 40$, $n \ne 11, 13$,
$$3.3 \le F_n \le 9.85\,,$$
2. $F_{11} = 12.1$, $F_{13} = 14.08$.

The value F_{11} is also achieved by a Barker sequence. From partial searches, for lengths up to 117, the highest known merit factor is between 8 and 9.56; for lengths from 118 to 200, the best-known factor is close to 6. For lengths > 200, statistical search methods have failed to yield a sequence having merit factor exceeding 5.

An *offset sequence* is one in which a fraction θ of the elements of a sequence of length n are chopped off at one end and appended to the other end, i.e., an offset sequence is a cyclic shift of the original sequence by $n\theta$ symbols. It turns out that the asymptotic merit factor of m sequences is equal to 3 and is independent of the particular offset of the m sequence. There exist offsets of sequences associated with quadratic-residue and twin-prime difference sets that achieve a larger merit factor of 6. Details may be found in [7].

8.5.3 Sequences with Low Aperiodic Crosscorrelation

If $\{u(t)\}$ and $\{v(t)\}$ are sequences of length $2n-1$ defined by

$$u(t) = \begin{cases} x(t) & \text{if } 0 \le t \le n-1 \\ 0 & \text{if } n \le t \le 2n-2 \end{cases}$$

and

$$v(t) = \begin{cases} y(t) & \text{if } 0 \le t \le n-1 \\ 0 & \text{if } n \le t \le 2n-2 \end{cases}$$

then

$$\{\rho_{x,y}(\tau) \mid -(n-1) \le \tau \le n-1\} = \left\{\theta_{u,v}(\tau) \mid 0 \le \tau \le 2n-2\right\} \tag{8.6}$$

Given a collection

$$U = \{\{x_i(t)\} \mid 1 \le i \le M\}$$

of sequences of length n over Z_q, let us define

$$\rho_{\max} = \max\left\{\left|\rho_{a,b}(\tau)\right| \mid a, b \in U\ ,\ \text{ either } a \ne b \text{ or } \tau \ne 0\right\}$$

It is clear from Eq. (8.6) how bounds on the *periodic* correlation parameter $\theta_{\max}$ can be adapted to give bounds on $\rho_{\max}$. Translation of the Welch bound gives that for every integer $k \ge 1$,

$$\rho_{\max}^{2k} \ge \left(\frac{n^{2k}}{M(2n-1)-1}\right)\left\{\frac{M(2n-1)}{\binom{2n+k-2}{k}} - 1\right\}$$

Setting $k = 1$ in the preceding bound gives

$$\rho_{\max} \ge n\sqrt{\frac{M-1}{M(2n-1)-1}}$$

Thus, for fixed M and large n, Welch's bound gives

$$\rho_{\max} \ge \mathcal{O}\left(n^{1/2}\right)$$

There exist sequence families which asymptotically achieve $\rho_{\max} \approx \mathcal{O}(n^{1/2})$, [10].

8.6 Other Correlation Measures

8.6.1 Partial-Period Correlation

The *partial-period* (*p-p*) *correlation* between the sequences $\{u(t)\}$ and $\{v(t)\}$ is the collection $\{\Delta_{u,v}(l,\tau,t_0) \mid 1 \le l \le n, 0 \le \tau \le n-1, 0 \le t_0 \le n-1\}$ of inner products

$$\Delta_{u,v}(l,\tau,t_0) = \sum_{t=t_0}^{t=t_0+l-1} u(t+\tau)v^*(t)$$

where l is the length of the partial period and the sum $t+\tau$ is again computed modulo n.

In direct-sequence CDMA systems, the pseudorandom signature sequences used by the various users are often very long for reasons of data security. In such situations, to minimize receiver hardware complexity, correlation over a partial period of the signature sequence is often used to demodulate data, as well as to achieve synchronization. For this reason, the p-p correlation properties of a sequence are of interest.

Researchers have attempted to determine the moments of the p-p correlation. Here the main tool is the application of the Pless power-moment identities of coding theory [8]. The identities often allow the first and second p-p correlation moments to be completely determined. For example, this is true in the case of m sequences (the remaining moments turn out to depend upon the specific characteristic polynomial of the m sequence). Further details may be found in [15].

8.6.2 Mean Square Correlation

Frequently in practice, there is a greater interest in the mean-square correlation distribution of a sequence family than in the parameter $\theta_{\max}$. Quite often in sequence design, the sequence family is derived from a linear, binary cyclic code of length n by picking a set of cyclically distinct sequences of period n. The families of Gold and Kasami sequences are so constructed. In this case, as pointed out by Massey, the mean square correlation of the family can be shown to be either optimum or close to optimum, under certain easily satisfied conditions, imposed on the minimum distance of the dual code. A similar situation holds even when the sequence family does not come from a linear cyclic code. In this sense, mean square correlation is not a very discriminating measure of the correlation properties of a family of sequences. An expanded discussion of this issue may be found in [5].

8.6.3 Optical Orthogonal Codes

Given a pair of $\{0, 1\}$ sequences $\{s_1(t)\}$ and $\{s_2(t)\}$ each having period n, we define the *Hamming correlation* function $\theta_{12}(\tau)$, $0 \leq \tau \leq n-1$, by

$$\theta_{12}(\tau) = \sum_{t=0}^{n-1} s_1(t+\tau)s_2(t)$$

Such correlations are of interest, for instance, in optical communication systems where the 1's and 0's in a sequence correspond to the presence or absence of pulses of transmitted light.

An (n, w, λ) optical orthogonal code (OOC) is a family $\mathcal{F} = \{\{s_i(t)\} \mid i = 1, 2, \ldots, M\}$, of M $\{0, 1\}$ sequences of period n, constant Hamming weight w, where w is an integer lying between 1 and $n-1$ satisfying $\theta_{ij}(\tau) \leq \lambda$ whenever either $i \neq j$ or $\tau \neq 0$.

Note that the Hamming distance $d_{a,b}$ between a period of the corresponding codewords $\{a(t)\}$, $\{b(t)\}$, $0 \leq t \leq n-1$ in an (n, w, λ) OOC having Hamming correlation ρ, $0 \leq \rho \leq \lambda$, is given by $d_{a,b} = 2(w - \rho)$, and, thus, OOCs are closely related to constant-weight error correcting codes. Given an (n, w, λ) OOC, by enlarging the OOC to include every cyclic shift of each sequence in the code, one obtains a constant-weight, minimum distance $d_{\min} \geq 2(w - \lambda)$ code. Conversely, given a constant-weight cyclic code of length n, weight w and minimum distance $d_{\min}$, one can derive an (n, w, λ) OOC code with $\lambda \leq w - d_{\min}/2$ by partitioning the code into cyclic equivalence classes and then picking precisely one representative from each equivalence class of size n.

By making use of this connection, one can derive bounds on the size of an OOC from known bounds on the size of constant-weight codes. The bound given next follows directly from the Johnson bound

for constant weight codes [8]. The number $M(n, w, \lambda)$ of codewords in a (n, w, λ) OOC satisfies

$$M(n, w, \lambda) \leq \frac{1}{w}\left\lfloor \frac{n-1}{w-1} \cdots \left\lfloor \frac{n-\lambda+1}{w-\lambda+1}\left\lfloor \frac{n-\lambda}{w-\lambda}\right\rfloor\right\rfloor \cdots \right\rfloor$$

An OOC code that achieves the Johnson bound is said to be optimal. A family $\{\mathcal{F}_n\}$ of OOCs indexed by the parameter n and arising from a common construction is said to be asymptotically optimum if

$$\lim_{n\to\infty} \frac{|\mathcal{F}_n|}{M(n, w, \lambda)} = 1$$

Constructions for optical orthogonal codes are available for the cases when $\lambda = 1$ and $\lambda = 2$. For larger values of λ, there exist constructions which are asymptotically optimum. Further details may be found in [6].

Defining Terms

Autocorrelation of a sequence: The complex inner product of the sequence with a shifted version itself.

Crosscorrelation of two sequences: The complex inner product of the first sequence with a shifted version of the second sequence.

***m* Sequence:** A periodic binary {0, 1} sequence that is generated by a shift register with linear feedback and which has maximal possible period given the number of stages in the shift register.

Pseudonoise sequences: Also referred to as pseudorandom sequences (PN), these are sequences that are deterministically generated and yet possess some properties that one would expect to find in randomly generated sequences.

Shift-register sequence: A sequence with symbols drawn from a field, which satisfies a linear-recurrence relation and which can be implemented using a shift register.

References

[1] Barg, A. On small families of sequences with low periodic correlation, *Lecture Notes in Computer Science*, 781, 154–158, Berlin, Springer-Verlag, 1994.

[2] Baumert, L.D. *Cyclic Difference Sets*, Lecture Notes in Mathematics 182, Springer–Verlag, New York, 1971.

[3] Boztaş, S., Hammons, R., and Kumar, P.V. 4-phase sequences with near-optimum correlation properties, *IEEE Trans. Inform. Theory*, IT-38, 1101–1113, 1992.

[4] Golomb, S.W. *Shift Register Sequences*, Aegean Park Press, San Francisco, CA, 1982.

[5] Hammons, A.R., Jr. and Kumar, P.V. On a recent 4-phase sequence design for CDMA. *IEICE Trans. Commun.*, E76-B(8), 1993.

[6] Helleseth, T. and Kumar, P.V. (planned). Sequences with low correlation. In *Handbook of Coding Theory*, ed., V.S. Pless and W.C. Huffman, Elsevier Science Publishers, Amsterdam, 1998.

[7] Jensen, J.M., Jensen, H.E., and Høholdt, T. The merit factor of binary sequences related to difference sets. *IEEE Trans. Inform. Theory*, IT-37(May), 617–626, 1991.

[8] MacWilliams, F.J. and Sloane, N.J.A. *The Theory of Error-Correcting Codes*, North-Holland, Amsterdam, 1977.

[9] Maschietti, A. Difference sets and hyperovals, *Designs, Codes and Cryptography*, 14, 89–98, 1998.
[10] Mow, W.H. On McEliece's open problem on minimax aperiodic correlation. In *Proc. IEEE Intern. Symp. Inform. Theory*, 75, 1994.
[11] Nechaev, A. The Kerdock code in a cyclic form, *Discrete Math. Appl.*, 1, 365–384, 1991.
[12] Peterson, W.W. and Weldon, E.J., Jr. *Error-Correcting Codes*, 2nd ed. MIT Press, Cambridge, MA, 1972.
[13] Sarwate, D.V. An upper bound on the aperiodic autocorrelation function for a maximal-length sequence. *IEEE Trans. Inform. Theory*, IT-30(July), 685–687, 1984.
[14] Sarwate, D.V. and Pursley, M.B. Crosscorrelation properties of pseudorandom and related sequences. *Proc. IEEE*, 68(May), 593–619, 1980.
[15] Simon, M.K., Omura, J.K., Scholtz, R.A., and Levitt, B.K. *Spread Spectrum Communications Handbook*, revised ed., McGraw Hill, New York, 1994.
[16] Solé, P. A quaternary cyclic code and a family of quadriphase sequences with low correlation properties, *Coding Theory and Applications*, *Lecture Notes in Computer Science*, 388, 193–201, Berlin, Springer-Verlag, 1989.
[17] Udaya, P. and Siddiqi, M. Optimal biphase sequences with large linear complexity derived from sequences over Z_4, *IEEE Trans. Inform. Theory*, IT-42 (Jan), 206–216, 1996.

Further Information

A more in-depth treatment of pseudonoise sequences, may be found in the following.

[1] Golomb, S.W. *Shift Register Sequences*, Aegean Park Press, San Francisco, 1982.
[2] Helleseth, T. and Kumar, P.V. Sequences with Low Correlation, in *Handbook of Coding Theory*, edited by V.S. Pless and W.C. Huffman, Elsevier Science Publishers, Amsterdam, 1998 (planned).
[3] Sarwate, D.V. and Pursley, M.B. Crosscorrelation Properties of Pseudorandom and Related Sequences, *Proc. IEEE*, 68, May, 593–619, 1980.
[4] Simon, M.K., Omura, J.K., Scholtz, R.A., and Levitt, B.K. *Spread Spectrum Communications Handbook*, revised ed., McGraw Hill, New York, 1994.