10.3.63 Remark The close relations between (n, w, λ) OOCs and constant weight codes provides good bounds on OOC from known bounds on constant weight codes. For further information see [1470].

10.3.64 Remark Other correlation measures include the *partial-period correlation* between two very long sequences where the correlation is calculated over a partial period. In practice, there is also some interest in the *mean-square correlation* of a sequence family rather than in $\theta_{\max}$. For the evaluation of these correlation measures coding theory sometimes plays an important role. For more information the reader is referred to [1470].

See Also

§6.1	Exponential sums are crucial in calculating the correlation of sequences.
§9.1	Boolean functions are closely related to sequences.
§9.3	Bent functions can be used in sequence constructions and vice versa.
§10.2	Linear Feedback Shift Registers (LFSRs) are important in constructing sequences with low correlation.
§14.6	Cyclic difference sets are related to sequences with two-level autocorrelation.
§17.3	Describes some applications of sequences and results on aperiodic correlation.

[1211]	An overview of some recent advances of low correlation sequences.
[1298]	This textbook by Golomb and Gong gives important information on sequences and their applications to signal design and cryptography.
[1470]	Provides an extensive survey of sequences with low correlation and their connections to coding theory.
[1471]	An elementary introduction to pseudonoise sequences.
[2518]	A classical paper on the crosscorrelation of pseudorandom sequences.

References Cited: [351, 381, 547, 701, 859, 1211, 1290, 1298, 1319, 1462, 1470, 1471, 1517, 1599, 1682, 1799, 1933, 2289, 2381, 2518, 2547, 2653, 2654, 2655, 2667, 2686, 2772, 2822, 2958]

10.4 Linear complexity of sequences and multisequences

Wilfried Meidl, *Sabanci University*
Arne Winterhof, *Austrian Academy of Sciences*

10.4.1 Linear complexity measures

10.4.1 Definition A sequence $S = s_0, s_1, \ldots$ over the finite field $\mathbb{F}_q$ is called a *(homogeneous) linear recurring sequence* over $\mathbb{F}_q$ with *characteristic polynomial*

$$f(x) = \sum_{i=0}^{l} c_i x^i \in \mathbb{F}_q[x]$$

of degree l, if S satisfies the *linear recurrence relation*

$$\sum_{i=0}^{l} c_i s_{n+i} = 0 \quad \text{for } n = 0, 1, \ldots. \tag{10.4.1}$$

10.4.2 Definition The *minimal polynomial* of a linear recurring sequence S is the uniquely defined monic polynomial $M \in \mathbb{F}_q[x]$ of smallest degree for which S is a linear recurring sequence with characteristic polynomial M. The *linear complexity* $L(S)$ of S is the degree of the minimal polynomial M.

10.4.3 Remark Without loss of generality one can assume that f is monic, i.e. $c_l = 1$. A sequence S over $\mathbb{F}_q$ is a linear recurring sequence if and only if S is ultimately periodic, if c_0 in (10.4.1) is nonzero then S is purely periodic, see [1933, Chapter 8]. Consequently Definition 10.4.1 is only meaningful for (ultimately) periodic sequences. Using the notation of [1130, 2058], we let $\mathcal{M}_q^{(1)}(f)$ be the set of sequences over $\mathbb{F}_q$ with characteristic polynomial f. The set of sequences with a fixed period N is then $\mathcal{M}_q^{(1)}(f)$ with $f(x) = x^N - 1$. The minimal polynomial M of a sequence $S \in \mathcal{M}_q^{(1)}(f)$ is always a divisor of f. For an N-periodic sequence S we have $L(S) \leq N$; see Section 10.2.

10.4.4 Remark The linear complexity of a sequence S can alternatively be defined as the length of the shortest linear recurrence relation satisfied by S. In engineering terms, $L(S)$ is 0 if S is the zero sequence and otherwise it is the length of the shortest linear feedback shift register (Section 10.2) that can generate S [1625, 1933, 2494, 2495].

10.4.5 Definition For $n \geq 1$ the *n-th linear complexity* $L(S, n)$ of a sequence S over $\mathbb{F}_q$ is the length L of a shortest linear recurrence relation

$$s_{j+L} = c_{L-1} s_{j+L-1} + \cdots + c_0 s_j, \quad 0 \leq j \leq n - L - 1,$$

over $\mathbb{F}_q$ satisfied by the first n terms of the sequence. The polynomial $\sum_{i=0}^{L} c_i x^i \in \mathbb{F}_q[x]$ is an *n-th minimal polynomial* of S. The *linear complexity* $L(S)$ of a periodic sequence can then be defined by

$$L(S) := \sup_{n \geq 1} L(S, n).$$

10.4.6 Remark Again one may assume that the n-th minimal polynomial is monic. Then it is unique whenever $L \leq n/2$. Definition 10.4.5 is also applicable for finite sequences, i.e. strings of elements of $\mathbb{F}_q$ of length n.

10.4.7 Definition For an infinite sequence S, the non-decreasing integer sequence $L(S, 1), L(S, 2), \ldots$ is the *linear complexity profile* of S.

10.4.8 Remark Linear complexity and linear complexity profile of a given sequence (as well as the linear recurrence defining it) can be determined by using the Berlekamp-Massey algorithm; see Section 15.1 or [1625, Section 6.7], and [2005]. The algorithm is efficient for sequences with low linear complexity and hence such sequences can easily be predicted.

10.4.9 Remark A sequence used as a keystream in stream ciphers must consequently have a large linear complexity, but also altering a few terms of the sequence should not cause a significant

decrease of the linear complexity. An introduction to the stability theory of stream ciphers is the monograph [868]. For a general comprehensive survey on the theory of stream ciphers we refer to [2494, 2495].

10.4.10 Definition The *k-error linear complexity* $L_k(S, n)$ of a sequence S of length n is defined by

$$L_k(S, n) = \min_T L(T, n),$$

where the minimum is taken over all sequences T of length n with Hamming distance $d(T, S)$ from S at most k. For an N-periodic sequence S over $\mathbb{F}_q$ the *k-error linear complexity* is defined by [2694]

$$L_k(S) = \min_T L(T),$$

where the minimum is taken over all N-periodic sequences T over $\mathbb{F}_q$ for which the first N terms differ in at most k positions from the corresponding terms of S.

10.4.11 Remark The concept of the k-error linear complexity is based on the *sphere complexity* introduced in [868].

10.4.12 Remark Recent developments in stream ciphers point toward an increasing interest in word-based or vectorized stream ciphers (see for example [779, 1440]), which requires the study of multisequences.

10.4.13 Definition For an arbitrary positive integer m, an *m-fold multisequence* $\mathbf{S} = (S_1, \ldots, S_m)$ over $\mathbb{F}_q$ (of finite or infinite length) is a string of m parallel sequences $S_1, \ldots, S_m$ over $\mathbb{F}_q$ (of finite or infinite length, respectively).

Let $f_1, \ldots, f_m \in \mathbb{F}_q[x]$ be arbitrary monic polynomials with $\deg(f_i) \geq 1$, $1 \leq i \leq m$. The set $\mathcal{M}_q(f_1, \ldots, f_m)$ is defined to be the set of m-fold multisequences $(S_1, \ldots, S_m)$ over $\mathbb{F}_q$ such that for each $1 \leq i \leq m$, S_i is a linear recurring sequence with characteristic polynomial f_i.

10.4.14 Definition The *joint minimal polynomial* of an m-fold multisequence $\mathbf{S} \in \mathcal{M}_q(f_1, \ldots, f_m)$ is the (uniquely determined) monic polynomial $M \in \mathbb{F}_q[x]$ of smallest degree which is a characteristic polynomial of S_i for all $1 \leq i \leq m$. The *joint linear complexity* of $\mathbf{S}$ is the degree of the joint minimal polynomial M.

10.4.15 Remark The set of N-periodic m-fold multisequences is $\mathcal{M}_q(f_1, \ldots, f_m)$ with $f_1 = \cdots = f_m = x^N - 1$, alternatively denoted by $\mathcal{M}_q^{(m)}(f)$ with $f(x) = x^N - 1$. The joint linear complexity of an m-fold multisequence can also be defined as the length of the shortest linear recurrence relation the m parallel sequences satisfy simultaneously. The joint minimal polynomial M of $\mathbf{S} \in \mathcal{M}_q(f_1, \ldots, f_m)$ is always a divisor of $\operatorname{lcm}(f_1, \ldots, f_m)$.

10.4.16 Definition For an integer $n \geq 1$ the *n-th joint linear complexity* $L(\mathbf{S}, n)$ of an m-fold multisequence $\mathbf{S} = (S_1, \ldots, S_m)$ is the length of the shortest linear recurrence relation the first n terms of the m parallel sequences $S_1, \ldots, S_m$ satisfy simultaneously. The *joint linear complexity profile* of $\mathbf{S}$ is the non-decreasing integer sequence $L(\mathbf{S}, 1), L(\mathbf{S}, 2), \ldots$.

10.4.17 Remark As the $\mathbb{F}_q$-linear spaces $\mathbb{F}_q^m$ and $\mathbb{F}_{q^m}$ are isomorphic, an m-fold multisequence $\mathbf{S}$ can also be identified with a single sequence $\mathcal{S}$ having its terms in the extension field $\mathbb{F}_{q^m}$. If $s_j^{(i)}$ denotes the j-th term of the i-th sequence S_i, $1 \leq i \leq m$, and $\{\beta_1, \ldots, \beta_m\}$ is a basis of

$\mathbb{F}_{q^m}$ over $\mathbb{F}_q$, then the j-th term of $\mathcal{S}$ is $\sigma_j = \sum_{i=1}^{m} \beta_i s_j^{(i)}$. The ($n$-th) joint linear complexity of $\mathcal{S}$ coincides then with the $\mathbb{F}_q$-*linear complexity* of $\mathcal{S}$, which is the length of the shortest linear recurrence relation with coefficients exclusively in $\mathbb{F}_q$ (the first n terms of) $\mathcal{S}$ satisfies (see [754, pp. 83–85]).

10.4.18 Definition We identify an m-fold multisequence $\mathbf{S}$ of length n (or period n) with an $m \times n$ matrix and write $\mathbf{S} \in \mathbb{F}_q^{m\times n}$ ($\mathbf{S} \in (\mathbb{F}_q^{m\times n})^\infty$). For two m-fold multisequences $\mathbf{S} = (S_1, \ldots, S_m), \mathbf{T} = (T_1, \ldots, T_m) \in \mathbb{F}_q^{m\times n}$ the *term distance* $d_T(\mathbf{S}, \mathbf{T})$ between $\mathbf{S}$ and $\mathbf{T}$ is the number of terms in the matrix for $\mathbf{S}$ that are different from the corresponding terms in the matrix for $\mathbf{T}$.

The *column distance* $d_C(\mathbf{S}, \mathbf{T})$ between $\mathbf{S}$ and $\mathbf{T}$ is the number of columns in which the matrices of $\mathbf{S}$ and $\mathbf{T}$ differ.

The *individual distances vector* for $\mathbf{S}, \mathbf{T}$ is defined by $d_V(\mathbf{S}, \mathbf{T}) = (d_H(S_1, T_1), \ldots, d_H(S_m, T_m))$, where d_H denotes the Hamming distance.

10.4.19 Example For $q = 2$, $m = 2$, $n = 5$, and

$$\mathbf{S} = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}, \quad \mathbf{T} = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix},$$

we have $d_T(\mathbf{S}, \mathbf{T}) = 3$, $d_C(\mathbf{S}, \mathbf{T}) = 2$ and $d_V(\mathbf{S}, \mathbf{T}) = (1, 2)$.

10.4.20 Definition For an integer k with $0 \le k \le mn$, the *(n-th) k-error joint linear complexity* $L_k(\mathbf{S}, n)$ of an m-fold multisequence $\mathbf{S}$ over $\mathbb{F}_q$ is defined by

$$L_k(\mathbf{S}, n) = \min_{\mathbf{T}\in\mathbb{F}_q^{m\times n}, d_T(\mathbf{S},\mathbf{T})\le k} L(\mathbf{T}, n).$$

For an integer $0 \le k \le n$ the *(n-th) k-error $\mathbb{F}_q$-linear complexity* $L_k^q(\mathbf{S}, n)$ of $\mathbf{S}$ is defined by

$$L_k^q(\mathbf{S}, n) = \min_{\mathbf{T}\in\mathbb{F}_q^{m\times n}, d_C(\mathbf{S},\mathbf{T})\le k} L(\mathbf{T}, n).$$

We define a partial order on $\mathbb{Z}^m$ by $\mathbf{k} = (k_1, \ldots, k_m) \le \mathbf{k}' = (k'_1, \ldots, k'_m)$ if $k_i \le k'_i$, $1 \le i \le m$. For $\mathbf{k} = (k_1, \ldots, k_m) \in \mathbb{Z}^m$ such that $0 \le k_i \le n$ for $1 \le i \le m$, the *(n-th) $\mathbf{k}$-error joint linear complexity* $L_{\mathbf{k}}(\mathbf{S}, n)$ of $\mathbf{S}$ is

$$L_{\mathbf{k}}(\mathbf{S}, n) = \min_{\mathbf{T}\in\mathbb{F}_q^{m\times n}, d_V(\mathbf{S},\mathbf{T})\le \mathbf{k}} L(\mathbf{T}, n),$$

i.e., the minimum is taken over all m-fold length n multisequences $\mathbf{T} = (T_1, \ldots, T_m)$ over $\mathbb{F}_q$ with Hamming distances $d_H(S_i, T_i) \le k_i$, $1 \le i \le m$.
The definitions for periodic multisequences are analogous.

10.4.2 Analysis of the linear complexity

10.4.21 Proposition Let $f \in \mathbb{F}_q[x]$ be a nonconstant monic polynomial.

1. [1625, Theorem 6.1.2], [1933, Chapter 8] For a sequence $S = s_0, s_1, \ldots$ over $\mathbb{F}_q$ consider the element $\sum_{i=0}^{\infty} s_i x^i$ in the ring $\mathbb{F}_q[[x]]$ of formal power series over $\mathbb{F}_q$. Then S is a linear recurring sequence with characteristic polynomial f if and only if $\sum_{i=0}^{\infty} s_i x^i = g(x)/f^*(x)$ with $g \in \mathbb{F}_q[x]$, $\deg(g) < \deg(f)$ and $f^*(x) = x^{\deg(f)} f(1/x)$ is the reciprocal polynomial of $f(x)$.

2. [2234, Lemma 1] For a sequence $S = s_1, s_2, \ldots$ over $\mathbb{F}_q$ consider the element $\sum_{i=1}^{\infty} s_i x^{-i}$ in the field $\mathbb{F}_q((x^{-1}))$ of formal Laurent series in x^{-1} over $\mathbb{F}_q$. Then S is a linear recurring sequence with characteristic polynomial f if and only if $\sum_{i=1}^{\infty} s_i x^{-i} = g(x)/f(x)$ with $g \in \mathbb{F}_q[x]$ and $\deg(g) < \deg(f)$.

10.4.22 Remark For more information and discussion of linear recurring sequences, we refer to Section 10.2.

10.4.23 Remark The reciprocal of a characteristic polynomial of a sequence S is also called a *feedback polynomial* of S.

10.4.24 Remark Proposition 10.4.21 implies a one-to-one correspondence between sequences in $\mathcal{M}_q^{(1)}(f)$ and rational functions g/f with $\deg(g) < \deg(f)$ (when the approach via Laurent series is used), and more generally between m-fold multisequences in $\mathcal{M}_q(f_1, \ldots, f_m)$ and m-tuples of rational functions $(g_1/f_1, \ldots, g_m/f_m)$ with $\deg(g_i) < \deg(f_i)$, $1 \le i \le m$. We note that in Proposition 10.4.21 Part 2 it is more convenient to start the indices for the sequence elements s_i with $i = 1$.

10.4.25 Proposition [1131] Let $(g_1/f_1, \ldots, g_m/f_m)$ be the m-tuple of rational functions corresponding to $\mathbf{S} \in \mathcal{M}_q(f_1, \ldots, f_m)$. The joint minimal polynomial of $\mathbf{S}$ is the unique monic polynomial $M \in \mathbb{F}_q[x]$ such that $\frac{g_1}{f_1} = \frac{h_1}{M}, \ldots, \frac{g_m}{f_m} = \frac{h_m}{M}$ for some (unique) polynomials $h_1, \ldots, h_m \in \mathbb{F}_q[x]$ with $\gcd(M, h_1, \ldots, h_m) = 1$.

10.4.26 Remark For an N-periodic sequence $S = s_0, s_1, \ldots$, let $S^N(x)$ be the polynomial $S^N(x) = s_0 + s_1 x + \cdots + s_{N-1} x^{N-1}$ of degree at most $N - 1$. Then $\sum_{i=0}^{\infty} s_i x^i = S^N(x)/(1 - x^N)$, which gives rise to the following theorem.

10.4.27 Theorem [754, Lemma 8.2.1], [2055] The joint linear complexity of an N-periodic m-fold multisequence $\mathbf{S} = (S_1, \ldots, S_m)$ is given by

$$L(\mathbf{S}) = N - \deg(\gcd(x^N - 1, S_1^N(x), \ldots, S_m^N(x))).$$

10.4.28 Remark Theorem 10.4.27 implies the famous Blahut theorem [302, 2495], [1625, Theorem 6.8.2] for the linear complexity of N-periodic sequences over $\mathbb{F}_q$, $\gcd(N, q) = 1$, which we state in 3 commonly used different versions.

10.4.29 Theorem (Blahut's Theorem) Let S be an N-periodic sequence over $\mathbb{F}_q$, let $\gcd(N, q) = 1$, and let α be a primitive N-th root of unity in an extension field of $\mathbb{F}_q$. Then

$$L(S) = N - |\{j \ : \ S^N(\alpha^j) = 0, 0 \le j \le N - 1\}|.$$

10.4.30 Theorem (Blahut's Theorem) Let $\gcd(N, q) = 1$, α be a primitive N-th root of unity in an extension field of $\mathbb{F}_q$ and let $A = (a_{ij})$ be the $N \times N$ Vandermonde matrix with $a_{ij} = \alpha^{ij}$, $0 \le i, j \le N - 1$. Let $\mathbf{s} = (s_0, s_1, \ldots, s_{N-1})$ be the vector corresponding to one period of an N-periodic sequence S over $\mathbb{F}_q$. The linear complexity $L(S)$ of S is the Hamming weight of the vector $A\mathbf{s}^T$.

10.4.31 Remark The vector $\mathbf{a} = A\mathbf{s}^T$ is called the *discrete Fourier transform* of $\mathbf{s}$. Several generalizations of the discrete Fourier transform have been suggested in the literature that can be used to determine the linear complexity of periodic sequences and multisequences with period not relatively prime to the characteristic of the field. We refer to [296, 2007, 2055].

10.4.32 Theorem (Blahut's Theorem) Let $S = s_0, s_1, \ldots$ be a sequence over $\mathbb{F}_q$ with period N dividing $q - 1$, and let $g \in \mathbb{F}_q[x]$ be the unique polynomial of degree at most $N - 1$ satisfying $g(\alpha^j) = s_j$, $j = 0, 1, \ldots$, where α is a fixed element of $\mathbb{F}_q$ of order N. Then $L(S) = w(g)$, where $w(g)$ denotes the weight of g, i.e., the number of nonzero coefficients of g.

10.4.33 Theorem [297, Theorem 8] Let f be a polynomial over a prime field $\mathbb{F}_p$ with degree of f at most $p-1$ and let $S = s_0, s_1, \ldots$ be the p-periodic sequence over $\mathbb{F}_p$ defined by $s_j = f(j)$, $j = 0, 1, \ldots$. Then $L(S) = \deg(f) + 1$.

10.4.34 Remark Theorem 8 in [297] more generally describes the linear complexity of p^r-periodic sequences over $\mathbb{F}_p$. A generalization of Theorem 10.4.33 to arbitrary finite fields is given in Theorem 1 of [2060].

10.4.35 Remark The linear complexity of an N-periodic sequence over $\mathbb{F}_q$ can be determined by the Berlekamp-Massey algorithm in $O(N^2)$ elementary operations. For some classes of period lengths, faster algorithms (of complexity $O(N)$) are known, the earliest being the Games-Chan algorithm [1165] for binary sequences with period $N = 2^v$. A collection of algorithms for several period lengths can be found in [3007] (see also [2952, 2953, 2954, 3008]). Some techniques to establish fast algorithms for arbitrary periods are presented in [85, 595, 596, 2051]. Stamp and Martin [2694] established a fast algorithm for the k-error linear complexity for binary sequences with period $N = 2^v$. Generalizations are presented in [1636, 1858, 2512], and for odd characteristic in [2050, 3007].

10.4.36 Remark In contrast to the faster algorithms introduced in the literature for certain period lengths, the Berlekamp-Massey algorithm also can determine the linear complexity profile of a (single) sequence. As an application, the general behaviour of linear complexity profiles can be analysed.

10.4.37 Theorem [1625, Theorem 6.7.4],[2494] Let $S = s_1, s_2, \ldots$ be a sequence over $\mathbb{F}_q$. If $L(S,n) > n/2$ then $L(S,n+1) = L(S,n)$. If $L(S,n) \leq n/2$, then $L(S,n+1) = L(S,n)$ for exactly one choice of $s_{n+1} \in \mathbb{F}_q$ and $L(S,n+1) = n+1-L(S,n)$ for the remaining $q-1$ choices of $s_{n+1} \in \mathbb{F}_q$.

10.4.38 Remark The linear complexity profile is uniquely described by the *increment sequence* of S, i.e., by the sequence of the positive integers among $L(S,1), L(S,2) - L(S,1), L(S,3) - L(S,2), \ldots$ [2240, 2929, 2931]. Another tool for the analysis of the linear complexity profile arises from a connection to the continued fraction expansion of Laurent series [2233, 2234].

10.4.39 Theorem [2234] Let $S = s_1, s_2, \ldots$ be a sequence over $\mathbb{F}_q$, let $S(x) = \sum_{i=1}^{\infty} s_i x^{-i} \in \mathbb{F}_q((x^{-1}))$ be the corresponding formal Laurent series, and let $A_1, A_2, \ldots$ be the polynomials in the continued fraction expansion of $S(x)$, i.e., $S(x) = 1/(A_1 + 1/(A_2 + \cdots))$ where $A_j \in \mathbb{F}_q[x]$, $\deg(A_j) \geq 1$, $j \geq 1$. Let $Q_{-1} = 0, Q_0 = 1$ and $Q_j = A_j Q_{j-1} + Q_{j-2}$ for $j \geq 1$. Then $L(S,n) = \deg(Q_j)$ where j is determined by

$$\deg(Q_{j-1}) + \deg(Q_j) \leq n < \deg(Q_j) + \deg(Q_{j+1}).$$

The n-th minimal polynomials are all (monic) polynomials of the form $M = aQ_j + gQ_{j-1}$, $a \in \mathbb{F}_q^*$, $g \in \mathbb{F}_q[x]$ with $\deg(g) \leq 2\deg(Q_j) - n - 1$. In particular, the increment sequence of S is $\deg(A_1), \deg(A_2), \ldots$.

10.4.40 Remark Generalizations of the Berlekamp-Massey algorithm and of continued fraction analysis for the linear complexity of multisequences can be found in [127, 758, 759, 760, 761, 868, 1049, 1665, 2508, 2509, 2938].

10.4.3 Average behaviour of the linear complexity

10.4.41 Remark We use the notation $N_n^{(m)}(L)$ and $E_n^{(m)}$ for the number of m-fold multisequences over $\mathbb{F}_q$ with length n and joint linear complexity L and the expected value for the joint linear complexity of a random m-fold multisequence over $\mathbb{F}_q$ of length n.

10.4.42 Theorem [1373, 2494, 2679] For $1 \le L \le n$

$$N_n^{(1)}(L) = (q-1)q^{\min(2L-1,2n-2L)}.$$

The expected value for $L(S,n)$ for a random sequence S over $\mathbb{F}_q$ is

$$E_n^{(1)} = \frac{1}{q^n} \sum_{S \in \mathbb{F}_q^n} L(S,n) = \begin{cases} \frac{n}{2} + \frac{q}{(q+1)^2} - q^{-n}\frac{n(q+1)+q}{(q+1)^2} & \text{for even } n, \\ \frac{n}{2} + \frac{q^2+1}{2(q+1)^2} - q^{-n}\frac{n(q+1)+q}{(q+1)^2} & \text{for odd } n. \end{cases}$$

10.4.43 Remark Theorem 10.4.42 was obtained by an analysis of the Berlekamp-Massey algorithm. Rueppel and Smeets [2494, 2679] provide closed formulas for the variance, showing that the variance is small. A detailed analysis of the linear complexity profile of sequences over $\mathbb{F}_q$ is given by Niederreiter in the series of papers [2229, 2233, 2234, 2237, 2240]. As a main tool, the continued fraction expansion of formal Laurent series is used. For a more elementary combinatorial approach, see [2236].

10.4.44 Theorem [2233] The linear complexity profile of a random sequence follows closely but irregularly the $n/2$-line, deviations from $n/2$ of the order of magnitude $\log n$ must appear for infinitely many n.

10.4.45 Remark The asymptotic behaviour of the joint linear complexity is investigated by Niederreiter and Wang in the series of papers [2269, 2270, 2927] using a sophisticated multisequence linear feedback shift-register synthesis algorithm based on a lattice basis reduction algorithm in function fields [2541, 2922, 2928].

10.4.46 Theorem [2247, 2269, 2270]

$$\begin{aligned} N_n^{(m)}(L) &= (q^m-1)q^{(m+1)L-m}, \quad 1 \le L \le n/2, \\ N_n^{(m)}(L) &\le C(q,m)L^m q^{2mn-(m+1)L}, \quad 1 \le L \le n, \end{aligned}$$

where $C(q,m)$ is a constant only depending on q and m. We have $N_n^{(m)}(L) \le q^{(m+1)L}$.

10.4.47 Remark In [2927] a method to determine $N_n^{(m)}(L)$ is presented and a closed formula for $N_n^{(2)}(L)$ is given. A closed formula for $N_n^{(3)}(L)$ is presented in [2270]. In [2269, 2270] it is shown that the joint linear complexity profile of a random m-fold multisequence follows closely the $mn/(m+1)$-line, generalizing Theorem 10.4.44 for $m = 1$.

10.4.48 Theorem [2269, 2270]

$$E_n^{(m)} = \frac{mn}{m+1} + o(n) \quad \text{as } n \to \infty.$$

For $m = 2, 3$ [1055, 2270, 2927]

$$E_n^{(m)} = \frac{mn}{m+1} + O(1), \quad \text{as } n \to \infty.$$

10.4.49 Remark Feng and Dai [1055] obtained their result with different methods, namely with multi-dimensional continued fractions.

10.4.50 Conjecture [2270]

$$E_n^{(m)} = \frac{mn}{m+1} + O(1) \quad \text{as } n \to \infty.$$

10.4.51 Remark For a detailed survey on recent developments in the theory of the n-th joint linear complexity of m-fold multisequences we refer to [2251].

10.4.52 Theorem [1130, 1132] For a monic polynomial $f \in \mathbb{F}_q[x]$ with $\deg(f) \geq 1$, let

$$f = r_1^{e_1} r_2^{e_2} \cdots r_k^{e_k}$$

be the canonical factorization of f into monic irreducible polynomials over $\mathbb{F}_q$. For $1 \leq i \leq k$, let $\alpha_i = q^{m \deg(r_i)}$. Then for an arbitrary positive integer m the expected value $E^{(m)}(f)$ of the joint linear complexity of a random m-fold multisequence from $\mathcal{M}_q^{(m)}(f)$ is

$$E^{(m)}(f) = \deg(f) - \sum_{i=1}^{k} \frac{1 - \alpha_i^{-e_i}}{\alpha_i - 1} \deg(r_i).$$

10.4.53 Remark In [1130, 1132] an explicit formula for the variance $\mathrm{Var}^{(m)}(f)$ of the joint linear complexity of random multisequences of $\mathcal{M}_q^{(m)}(f)$ is given. In [1131, 1132] it is shown how to obtain from Theorem 10.4.52 closed formulas for the more general case of m-fold multisequences in $\mathcal{M}_q(f_1, \ldots, f_m)$.

10.4.54 Remark Since for $f(x) = x^N - 1$ the set $\mathcal{M}^{(m)}(f)$ is the set of N-periodic sequences, earlier formulas on expectation (and variance) of the (joint) linear complexity of periodic (multi)sequences can be obtained as a corollary of Theorem 10.4.52: [2053, Theorem 3.2], [2054, Theorem 1], [3019, Theorem 1] on $E^{(1)}(x^N - 1)$, and [1133, Theorem 1], [2055, Theorem 1] on $E^{(m)}(x^N - 1)$ for arbitrary m.

10.4.55 Remark In [1133, 2055] lower bounds on the expected joint linear complexity for periodic multisequences are presented, estimating the magnitude of the formula for $E^{(m)}(x^N - 1)$ in Theorem 10.4.52. In [1133] it is also noted that the variance $\mathrm{Var}^{(m)}(x^N - 1)$ is small, showing that for random N-periodic multisequences over $\mathbb{F}_q$ the joint linear complexity is close to N (the trivial upper bound), with a small variance.

10.4.56 Remark Lower bounds for the expected n-th k-error joint linear complexity, the expected n-th k-error $\mathbb{F}_q$-linear complexity and the expected n-th $\mathbf{k}$-error joint linear complexity for an integer vector $\mathbf{k} = (k_1, \ldots, k_m)$ for a random m-fold multisequence over $\mathbb{F}_q$ are established in [2057]. These results generalize earlier bounds for the case $m = 1$ presented in [2052].

10.4.57 Remark For periodic sequences, lower bounds on the expected k-error linear complexity have been established in [2053, 2054]. For periodic multisequences (with prime period N different from the characteristic), lower bounds for the expected error linear complexity are presented in [2057] for all 3 multisequence error linear complexity measures.

10.4.58 Remark In the papers [2056, 2248, 2267, 2268, 2860] the question is addressed if linear complexity and k-error linear complexity can be large simultaneously. Among others, the existence of N-periodic sequences attaining the upper bounds N and $N - 1$ for linear and k-error linear complexity is shown for infinitely many period lengths (and a certain range for k depending on the period length), and it is shown that for several classes of period length a large number of N-periodic (multi)sequences with (joint) linear complexity N also exhibits a large k-error linear complexity.

10.4.59 Remark In [3010] methods from function fields are used to construct periodic multisequences with large linear complexity and k-error linear complexity simultaneously for various period lengths.

10.4.4 Some sequences with large n-th linear complexity

10.4.4.1 Explicit sequences

10.4.60 Definition For $a, b \in \mathbb{F}_p$ with $a \neq 0$ the *explicit inversive congruential sequence* $Z = z_0, z_1, \ldots$ is

$$z_j = (aj + b)^{p-2}, \quad j \geq 0. \tag{10.4.2}$$

10.4.61 Theorem [2061] We have

$$L(Z, n) \geq \begin{cases} (n-1)/3 & \text{for } 1 \leq n \leq (3p-7)/2, \\ n - p + 2 & \text{for } (3p-5)/2 \leq n \leq 2p-3, \\ p - 1 & \text{for } n \geq 2p-2. \end{cases}$$

10.4.62 Remark We note that $j^{p-2} = j^{-1}$ for $j \in \mathbb{F}_p^*$. Since inversion is a fast operation this sequence is despite its high n-th linear complexity still highly predictable.

10.4.63 Remark Analogous sequences of (10.4.2) over arbitrary finite fields $\mathbb{F}_q$ are studied in [2061]. Multisequences of this form are investigated in [2064]. Explicit inversive sequences and multisequences can also be defined using the multiplicative structure of $\mathbb{F}_q$.

10.4.64 Definition For $m \geq 1$, $\alpha_i, \beta_i \in \mathbb{F}_q^*$, $1 \leq i \leq m$, and an element $\gamma \in \mathbb{F}_q$ of order N, the *explicit inversive congruential sequence of period* N, $\mathbf{Z} = (Z_1, \ldots, Z_m)$, with $Z_i = \sigma_0^{(i)}, \sigma_1^{(i)}, \ldots$ is

$$\sigma_j^{(i)} = (\alpha_i \gamma^j + \beta_i)^{q-2}, \quad j \geq 0. \tag{10.4.3}$$

10.4.65 Remark Sequences of the form (10.4.3) are analysed in [2063, 2064]. With an appropriate choice of the parameters one can obtain (multi)sequences with *perfect linear complexity profile*, i.e., $L(\mathbf{Z}, n) \geq mn/(m+1)$.

10.4.66 Theorem [2064] Let $m < (q-1)/N$ and let $C_1, \ldots, C_m$ be different cosets of the group $\langle \gamma \rangle$ generated by γ, such that none of them contains the element -1. For $1 \leq i \leq m$ choose α_i, β_i such that $\alpha_i \beta_i^{-1} \in C_i$, then

$$L(\mathbf{Z}, n) \geq \min\left\{ \frac{mn}{m+1}, N \right\}, \quad n \geq 1.$$

10.4.67 Definition Given an element $\vartheta \in \mathbb{F}_q^*$, the *quadratic exponential sequence* $Q = q_0, q_1, \ldots$ is

$$q_j = \vartheta^{j^2}, \qquad j \geq 0.$$

10.4.68 Theorem [1377] We have

$$L(Q, n) \geq \frac{\min\{n, N\}}{2}, \quad n \geq 1.$$

10.4.69 Remark The period N of Q is at least half of the multiplicative order of ϑ.

10.4.4.2 Recursive nonlinear sequences

10.4.70 Definition Given a polynomial $f \in \mathbb{F}_p[x]$ of degree $d \geq 2$, the *nonlinear congruential sequence* $U = u_0, u_1, \ldots$ is defined by the recurrence relation

$$u_{j+1} = f(u_j), \quad j \geq 0, \tag{10.4.4}$$

with some initial value $u_0 \in \mathbb{F}_p$ such that U is purely periodic with some period $N \leq p$.

10.4.71 Theorem [1377] Let U be as in (15.1.8), where $f \in \mathbb{F}_p[x]$ is of degree $d \geq 2$, then

$$L(U, n) \geq \min\left\{\log_d(n - \lfloor \log_d n \rfloor), \log_d N\right\}, \quad n \geq 1.$$

10.4.72 Remark For some special classes of polynomials much better results are available, see [1354, 1377, 2636]. For instance, in case of the largest possible period $N = p$ we have

$$L(U, n) \geq \min\{n - p + 1, p/d\}, \quad n \geq 1.$$

10.4.73 Theorem [1377] The *inversive (congruential) sequence* $Y = y_0, y_1, \ldots$ defined by

$$y_{j+1} = a y_j^{p-2} + b, \quad j \geq 0,$$

with $a, b, y_0 \in \mathbb{F}_p$, $a \neq 0$, has linear complexity profile

$$L(Y, n) \geq \min\left\{\frac{n-1}{3}, \frac{N-1}{2}\right\}, \quad n \geq 1.$$

10.4.74 Theorem [1354, 2636] The *power sequence* $P = p_0, p_1, \ldots$, defined as

$$p_{j+1} = p_j^e, \quad j \geq 0,$$

with some integer $e \geq 2$ and initial value $0 \neq p_0 \in \mathbb{F}_p$ satisfies

$$L(P, n) \geq \min\left\{\frac{n^2}{4(p-1)}, \frac{N^2}{p-1}\right\}, \quad n \geq 1.$$

10.4.75 Remark Two more classes of nonlinear sequences provide much better results than in the general case, nonlinear sequences with *Dickson polynomials* [87] and *Rédei functions* [2066]. See Section 9.6 and [1930] for the definitions.

10.4.4.3 Legendre sequence and related bit sequences

10.4.76 Definition Let $p > 2$ be a prime. The *Legendre sequence* $\Lambda = l_0, l_1, \ldots$, for $j \geq 0$, is

$$l_j = \begin{cases} 1 & \text{if } \left(\frac{j}{p}\right) = -1, \\ 0 & \text{otherwise,} \end{cases}$$

where $\left(\frac{\cdot}{p}\right)$ is the Legendre symbol.

10.4.77 Theorem [754, 2823] The linear complexity of the Legendre sequence is

$$L(\Lambda) = \begin{cases} (p-1)/2 & \text{if } p \equiv 1 \pmod 8, \\ p & \text{if } p \equiv 3 \pmod 8, \\ p-1 & \text{if } p \equiv 5 \pmod 8, \\ (p+1)/2 & \text{if } p \equiv 7 \pmod 8. \end{cases}$$

10.4.78 Theorem [2638, Theorem 9.2] The linear complexity profile of the Legendre sequence satisfies

$$L(\Lambda, n) > \frac{\min\{n, p\}}{1 + p^{1/2}(1 + \log p)} - 1, \quad n \geq 1.$$

10.4.79 Remark For similar sequences, that are defined by the use of the quadratic character of arbitrary finite fields and the study of their linear complexity profiles, see [1780, 2059, 2988].

10.4.80 Definition Let γ be a primitive element and η be the quadratic character of the finite field $\mathbb{F}_q$ of odd characteristic. The *Sidelnikov sequence* $\sigma = \sigma_0, \sigma_1, \ldots$ for $j \geq 0$, is

$$\sigma_j = \begin{cases} 1 & \text{if } \eta(\gamma^j + 1) = -1, \\ 0 & \text{otherwise.} \end{cases}$$

10.4.81 Remark In many cases one is able to determine the linear complexity $L(\sigma)$ over $\mathbb{F}_2$ exactly, see Meidl and Winterhof [2065]. For example, if $(q-1)/2$ is an odd prime such that 2 is a primitive root modulo $(q-1)/2$, then σ attains the largest possible linear complexity $L(\sigma) = q - 1$. Moreover we have the lower bound [2065]

$$L(\sigma, n) \gg \frac{\min\{n, q\}}{q^{1/2} \log q}, \quad n \geq 1.$$

The k-error linear complexity of the Sidelnikov sequence seen as a sequence over $\mathbb{F}_p$ has been estimated in [86, 636, 1194]. For results on similar sequences with composite modulus see [390] and [754, Chapter 8.2].

10.4.4.4 Elliptic curve sequences

10.4.82 Definition Let $p > 3$ be a prime and E be an elliptic curve over $\mathbb{F}_p$ of the form

$$Y^2 = X^3 + aX + b$$

with coefficients $a, b \in \mathbb{F}_p$ such that $4a^3 + 27b^2 \neq 0$. For a given initial point $W_0 \in E(\mathbb{F}_p)$, a fixed point $G \in E(\mathbb{F}_p)$ of order N and a rational function $f \in \mathbb{F}_p(E)$ the *elliptic curve congruential sequence* $W = w_0, w_1, \ldots$ (with respect to f) is

$$w_j = f(W_j), \quad j \geq 0, \quad \text{where} \quad W_j = G \oplus W_{j-1} = jG \oplus W_0, \quad j \geq 1.$$

10.4.83 Remark Obviously, W is N-periodic.

10.4.84 Remark For example, choosing the function $f(x, y) = x$, the work of Hess and Shparlinski [1488] gives the lower bound

$$L(W, n) \geq \min\{n/3, N/2\}, \quad n \geq 2.$$

10.4.5 Related measures

10.4.5.1 Kolmogorov complexity

10.4.85 Remark The *Kolmogorov complexity* is a central topic in *algorithmic information theory*. The Kolmogorov complexity of a binary sequence is, roughly speaking, the length of the

shortest computer program that generates the sequence. The relationship between linear complexity and Kolmogorov complexity was studied in [256, 2940]. The Kolmogorov complexity is twice the linear complexity for almost all sequences over $\mathbb{F}_2$ of sufficiently (but only moderately) large length. In contrast to the linear complexity the Kolmogorov complexity is in general not computable and so of no practical significance.

10.4.5.2 Lattice test

10.4.86 Definition Let $S = s_0, s_1, \ldots$ be a sequence over $\mathbb{F}_q$, and for $s \geq 1$ let $V(S, s)$ be the subspace of $\mathbb{F}_q^s$ spanned by the vectors $\mathbf{s}_j - \mathbf{s}_0$, $j = 1, 2, \ldots$, where

$$\mathbf{s}_j = (s_j, s_{j+1}, \ldots, s_{j+s-1}), \quad j \geq 0.$$

The sequence S passes the *s-dimensional lattice test* for some $s \geq 1$, if $V(S, s) = \mathbb{F}_q^s$. For given $s \geq 1$ and $n \geq 2$ we say that S passes the *s-dimensional n-lattice test* if the subspace spanned by the vectors $\mathbf{s}_j - \mathbf{s}_0$, $1 \leq j \leq n - s$, is $\mathbb{F}_q^s$. The largest s for which S passes the s-dimensional n-lattice test is the *lattice profile at n* and is denoted by $\mathcal{S}(S, n)$.

10.4.87 Theorem [909] We have either

$$\begin{aligned} \mathcal{S}(S,n) &= \min\{L(S,n), n+1-L(S,n)\} \quad \text{or} \\ \mathcal{S}(S,n) &= \min\{L(S,n), n+1-L(S,n)\} - 1. \end{aligned}$$

10.4.88 Remark The results of [908] on the expected value of the lattice profile show that a "random" sequence should have $\mathcal{S}(S, n)$ close to $\min\{n/2, N\}$.

10.4.5.3 Correlation measure of order k

10.4.89 Definition The *correlation measure of order k* of a binary sequence S is

$$C_k(S) = \max_{M,D} \left| \sum_{n=0}^{M-1} (-1)^{s_{n+d_1}} \cdots (-1)^{s_{n+d_k}} \right|, \quad k \geq 1,$$

where the maximum is taken over all $D = (d_1, d_2, \ldots, d_k)$ with non-negative integers $d_1 < d_2 < \cdots < d_k$ and M such that $M - 1 + d_k \leq T - 1$. Obviously, $C_2(S)$ is bounded by the maximal absolute value of the aperiodic autocorrelation of S.

10.4.90 Remark The correlation measure of order k was introduced by Mauduit and Sárközy in [2031]. The linear complexity profile of a given N-periodic sequence can be estimated in terms of its correlation measure and a lower bound on $L(S, n)$ can be obtained whenever an appropriate bound on $\max C_k(S)$ is known.

10.4.91 Theorem [391] We have

$$L(S,n) \geq n - \max_{1 \leq k \leq L(S,n)+1} C_k(S), \quad 1 \leq n \leq N-1.$$

10.4.5.4 FCSR and p-adic span

10.4.92 Remark In [1742] an alternative feedback shift register architecture was presented, *feedback with carry shift registers (FCSR)*. For binary sequences the procedure is as follows: Differently to linear recurring sequences the bits are added as integers (again following a linear

recurrence relation). The result is added to the content of a memory, which is a nonnegative integer m, to obtain an integer σ. The parity bit $\sigma \pmod 2$, of σ is then the next term of the sequence, and the higher order bits $\lfloor \sigma/2 \rfloor$ are the new content of the memory.

FCSR-sequences share many properties with linear recurring sequences, but for their analysis instead of arithmetics in finite fields, arithmetics in the 2-adic numbers is used - or in the more general case of sequences modulo p in the p-adic numbers.

An FCSR-equivalent to the linear complexity is the 2-*adic span*, respectively the *p-adic span* of a sequence, which measures the size of the smallest FCSR that generates the sequence.

Since their introduction, FCSR-sequences attracted a lot of attention. We refer to [129, 1325, 1326, 1743, 2768] and the references therein.

10.4.5.5 Discrepancy

10.4.93 Definition Let $X = x_0, x_1, \ldots$ be a sequence in the unit interval $[0, 1)$. For $0 \le d_1 < \cdots < d_k < n$ we put

$$\mathbf{x}_j = \mathbf{x}_j(d_1, \ldots, d_k) = (x_{j+d_1}, \ldots, x_{j+d_k}), \quad 1 \le j \le n - d_k.$$

The *discrepancy* of the vectors $\mathbf{x}_1(d_1, \ldots, d_k), \ldots, \mathbf{x}_{n-d_k}(d_1, \ldots, d_k)$ is

$$\sup_I \left| \frac{A(I, \mathbf{x}_1, \ldots, \mathbf{x}_{n-d_k})}{n - d_k} - V(I) \right|,$$

where the supremum is taken over all subintervals of $[0, 1)^k$, $V(I)$ is the volume of I and $A(I, \mathbf{x}_1, \ldots, \mathbf{x}_{n-d_k})$ is the number of points $\mathbf{x}_j$, $j = 1, \ldots, n - d_k$, in the interval I.

10.4.94 Remark We can derive a binary sequence $B = e_0, e_1, \ldots$ from X by $e_j = 1$ if $0 \le x_j < 1/2$ and $e_j = 0$ otherwise.

10.4.95 Remark In [2030, Theorem 1] the correlation measure of order k of B is estimated in terms of the above discrepancy of vectors derived from the sequence X. Hence, using the relation between linear complexity profile and correlation measure of B we can obtain (weak) linear complexity profile lower bounds for B from discrepancy upper bounds for X.

See Also

§6.3, §10.2, §17.3	For related measures.
§9.1, §9.3	For Boolean functions and nonlinearity.
§10.2	For LFSR.
§10.5, §17.1	For nonlinear recurrence sequences.
§12.2, §12.3, §16.4	For elliptic curves.
§15.1	For basics on coding theory and the Berlekamp-Massey algorithm.
§16.2	For stream ciphers.

References Cited: [85, 86, 87, 127, 129, 256, 296, 297, 302, 390, 391, 595, 596, 636, 754, 758, 759, 760, 761, 779, 868, 908, 909, 1049, 1055, 1130, 1131, 1132, 1133, 1165, 1194, 1325, 1326, 1354, 1373, 1377, 1440, 1488, 1625, 1636, 1665, 1742, 1743, 1780, 1858, 1930, 1933, 2005, 2007, 2030, 2031, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2063, 2064, 2065, 2066, 2229, 2233, 2234, 2236, 2237, 2240, 2247, 2248, 2251, 2267, 2268,