Golay Complementary Sequences

Matthew G. Parker*, Kenneth G. Paterson† and Chintha Tellambura‡

June 22, 2004

1 Introduction

Complementary sequences were introduced by Marcel Golay [1] in the context of infrared spectrometry. A *complementary pair* of sequences (CS pair) satisfies the useful property that their out-of-phase aperiodic autocorrelation coefficients sum to zero [1, 2]. Let $\mathbf{a} = (a_0, a_1, \ldots, a_{N-1})$ be a sequence of length N such that $a_i \in \{+1, -1\}$ (we say that $\mathbf{a}$ is bi-polar). Define the Aperiodic Auto-Correlation Function (AACF) of $\mathbf{a}$ by

$$\rho_{\mathbf{a}}(k) = \sum_{i=0}^{N-k-1} a_i a_{i+k}, \qquad 0 \le k \le N-1. \tag{1}$$

Let $\mathbf{b}$ be defined similarly to $\mathbf{a}$. The pair $(\mathbf{a}, \mathbf{b})$ is called a Golay Complementary Pair (GCP) if:

$$\rho_{\mathbf{a}}(k) + \rho_{\mathbf{b}}(k) = 0, \qquad k \neq 0. \tag{2}$$

Each member of a GCP is called a Golay complementary sequence (GCS, or simply Golay sequence). Note that this definition (2) can be generalised to non-binary sequences. For instance, a_i and b_i can be selected from the set $\{\zeta^0, \zeta^1, \ldots, \zeta^{2^h-1}\}$ where ζ is a primitive q-th root of unity, which yields so-called polyphase Golay sequences. In this survey we however emphasise binary GCPs.

*M.G.Parker is with the Code Theory Group, Inst. for Informatikk, Høyteknologisenteret i Bergen, University of Bergen, Bergen 5020, Norway. E-mail: matthew@ii.uib.no. Home Page: http://www.ii.uib.no/~matthew/, Author funded by NFR Project Number 119390/431

†K.G.Paterson is with the Information Security Group, Royal Holloway, University of London, Egham, Surrey TW20 0EX, U.K. E-mail:kenny.paterson@rhul.ac.uk. Author funded by the Nuffield Foundation, NUF-NAL 02.

‡C. Tellambura is with the Department of Electrical and Computer Engineering, Electrical and Computer Engineering Research Facility, University of Alberta, Edmonton, Alberta, CANADA, T6G 2V4. E-mail: chintha@ee.ualberta.ca. Phone/Fax: +780-492-7228(1811)

It is helpful to view (2) in polynomial form. A sequence $\mathbf{a}$ can be associated with the polynomial $a(z) = a_{N-1}z^{N-1} + a_{N-2}z^{N-2} + \ldots + a_1 z + a_0$ in indeterminate z with coefficients ± 1. The pair $(\mathbf{a}, \mathbf{b})$ is then a GCP if the associated polynomials $(a(z), b(z))$ satisfy

$$a(z)a(z^{-1}) + b(z)b(z^{-1}) = 2N. \tag{3}$$

Equations (2) and (3) are equivalent expressions because
$a(z)a(z^{-1}) = \rho_{\mathbf{a}}(0) + \sum_{1}^{N-1} \rho_{\mathbf{a}}(k)(z^k + z^{-k})$. A further condition can be obtained by restricting z to lie on the unit circle in the complex plane, i.e. $z \in \{e^{2\pi jt} | j^2 = -1, 0 \leq t < 1\}$. Then $|a(z)|^2 = a(z)a(z^{-1})$ and we have

$$|a(z)|^2 + |b(z)|^2 = 2N, \quad |z| = 1. \tag{4}$$

This means that the absolute value of each polynomial on the unit circle is bounded by $\sqrt{2N}$.

Golay complementary pairs and sequences have found application in physics (Ising spin systems), combinatorics (orthogonal designs and Hadamard matrices) and telecommunications (for instance, to surface-acoustic wave design, the Loran C precision navigation system, channel-measurement, optical time-domain reflectometry [3], synchronisation, spread-spectrum communications, and, recently, Orthogonal Frequency Division Multiplexing (OFDM) systems [4, 5, 6, 7]. Initially, the properties of the *pair* were primarily exploited [1] in a two-channel setting and Periodic GCPs have lately been proposed for two-sided channel-estimation, where the two sequences in the pair form a pre-amble and post-amble training sequence, respectively [8]. In recent years, the spectral spread properties of each individual sequence in the pair have also been used. As just one example of this, we briefly describe the application of Golay sequences in OFDM. Here, given a data sequence, $\mathbf{a} = (a_0, a_1, \ldots, a_{N-1})$, the transmitted signal $s_{\mathbf{a}}(t)$ as a function of time t is essentially the real part of a Discrete Fourier Transform (DFT) of $\mathbf{a}$:

$$s_{\mathbf{a}}(t) = \sum_{i=0}^{N-1} a_i e^{2\pi j(i\Delta f + f_0)t} \tag{5}$$

where Δf the frequency separation between adjacent subcarrier pairs and f_0 is the base frequency. Notice that $|s_{\mathbf{a}}(t)| = |a(e^{2\pi j i\Delta f t})|$ where $a(z)$ is the polynomial corresponding to $\mathbf{a}$. Thus the power characteristics of the OFDM signal can be studied by examining the behaviour of an associated polynomial on $|z| = 1$. In particular, if $\mathbf{a}$ is a GCS, then we have that $|s_{\mathbf{a}}(t)|^2 \leq 2N$ so that the Peak-to-Mean Envelope Power Ratio (PMEPR) of the signal is at most 2.0. Having such tightly-bounded OFDM signals eases amplifier specification at the OFDM transmitter.

Let $\mathbf{A} = (A_0, A_1, \ldots, A_{N'-1})$ be the N'-point oversampled DFT of $\mathbf{a}$, where $N' \geq N$, i.e.

$$A_k = \sum_{i=0}^{N-1} a_i \omega^{ik} = a(\omega^k) \quad 0 \leq k < N',$$

where $\omega = e^{2\pi j/N'}$ is a complex N'-th root of unity. For N' large, the values of the N'-point oversampled DFT of $\mathbf{a}$ can be used to approximate the values $a(z)$, $|z| = 1$, and thus the complex OFDM signal in (5).

Example 1: Let $\mathbf{a} = - + + - + - + + + -$, $\mathbf{b} = - + + + + + + - - +$, where '$+$' and '$-$' mean 1 and -1, respectively. The AACFs of $\mathbf{a}$ and $\mathbf{b}$ are:

$$\rho_{\mathbf{a}}(k) = (10, -3, 0, -1, 0, 1, 2, -1, -2, 1),$$
$$\rho_{\mathbf{b}}(k) = (10, 3, 0, 1, 0, -1, -2, 1, 2, -1).$$

It is evident that the AACFs of $\mathbf{a}$ and $\mathbf{b}$ sum to a δ-function, as required by (2) and (3), so $(\mathbf{a}, \mathbf{b})$ is a GCP. The absolute squared values of the 20-point oversampled DFT of $\mathbf{a}$ and $\mathbf{b}$ are:

$$\mathbf{A} = 10 \cdot (0.40, 0.44, 0.15, 0.73, 1.85, 0.20, 1.05, 1.67, 0.95, 1.96, 1.60, 1.96, 0.95, 1.67, 1.05, 0.20, 1.85, 0.73, 0.15, 0.44)$$
$$\mathbf{B} = 10 \cdot (1.60, 1.56, 1.85, 1.27, 0.15, 1.80, 0.95, 0.33, 1.05, 0.04, 0.40, 0.04, 1.05, 0.33, 0.95, 1.80, 0.15, 1.27, 1.85, 1.56)$$

At every point these two power spectra add to 20, as required by (4).

It should be stressed that the bound of $\sqrt{2N}$ on the amplitude of $a(z)$ on $|z| = 1$ is extremely low for any bipolar sequence of length greater or equal to about 16. One would not find such sequences by chance. And the complementary sequence/aperiodic correlation approach is currently the only construction method known that tightly upper bounds these values for bi-polar sequences. There is also the Rudin-Shapiro (RuS) construction [9], which appeared soon after Golay's initial work, but the RuS construction can be viewed as a basic recursive Golay construction technique. This construction is described in Section 2.2. Indeed research on the uniformity of polynomials on the unit circle has continued largely independently in the mathematical community for many years [10, 11, 12, 13], and this work indicates that sequences with good AACFs and flat DFT spectra, or equivalently, polynomials that are approximately uniform on $|z| = 1$, are rather difficult to construct. For example, the celebrated conjecture of Littlewood on flat polynomials on the unit circle is still open:

Conjecture 1 *[12] There exists a pair of constants C_0, C_1 and a series of degree $N-1$ polynomials $a(z)$ with ± 1 coefficients such that, as $N \to \infty$,*

$$C_0\sqrt{N} \leq |a(z)| \leq C_1\sqrt{N}, \quad |z| = 1. \tag{6}$$

There is no known construction that produces polynomials satisfying the lower bound of Conjecture 1, and the complementary sequence approach is the only one known that gives polynomials satisfying the upper bound.

2 Existence and Construction

2.1 Necessary Conditions

As we shall see in the next section, GCPs are known to exist for all lengths $N = 2^\alpha 10^\beta 26^\gamma$, $\alpha, \beta, \gamma \geq 0$, [14]. GCPs are not known for any other lengths. Golay showed that the length N of a Golay sequence must be the sum of two squares (where one square may be 0) [2]. More recently it has been shown that GCPs of length N do not exist if there is a prime p with $p = 3(\mathrm{mod} 4)$ such that $p|N$, [16]. This generalised earlier, weaker non-existence results. Therefore, the admissable lengths < 100 are:

$$1, 2, 4, 8, 10, 16, 20, 26, 32, 34^*, 40, 50^*, 52, 58^*, 64, 68^*, 74^*, 80, 82^*.$$

Moreover, various computer searches have eliminated lengths marked by '$*$' in the above list. Therefore the lengths, $N < 100$, for which GCPs exist are:

$$1, 2, 4, 8, 10, 16, 20, 26, 32, 40, 52, 64, 80 \tag{7}$$

In the next sections, we provide constructions covering all these lengths.

2.2 Recursive Constructions

Using (3), many recursive constructions for GCPs can be obtained via simple algebraic manipulation. For instance, if $a(z)$ and $b(z)$ are a Golay pair of length N, then simple algebraic manipulation shows that $a(z) + z^N b(z)$ and $a(z) - z^N b(z)$ also satisfy (3) with $2N$ being replaced by $4N$. This is in fact the well-known Golay-Rudin-Shapiro recursion, generating a length $2N$ GCP from a length N GCP [13]. We may write this more simply in terms of sequences as:

$$(\mathbf{a}, \mathbf{b}) \rightarrow (\mathbf{a}|\mathbf{b}, \quad \mathbf{a}|\overline{\mathbf{b}}) \tag{8}$$

where '$|$' means concatenation.

The following are a few other recursive constructions:

- The construction of Turyn [14] can be stated as follows. Let $(\mathbf{a}, \mathbf{b})$ and $(\mathbf{c}, \mathbf{d})$ be GCPs of length M and N, respectively. Then

$$\begin{aligned} &a(z^N)(c(z) + d(z))/2 + z^{N(M-1)} b(z^{-N})(c(z) - d(z))/2, \\ &b(z^N)(c(z) + d(z))/2 - z^{N(M-1)} a(z^{-N})(c(z) - d(z))/2 \end{aligned} \tag{9}$$

  is a GCP of length MN.

- The constructions of Golay in [2] are obtained as follows. Let $(\mathbf{a}, \tilde{\mathbf{b}})$ and $(\mathbf{c}, \mathbf{d})$ be GCPs of lengths M and N, respectively, where $\tilde{\mathbf{b}}$ means reversal of $\mathbf{b}$.

  Golay's *concatenation* construction can be stated as:

$$a(z^N)c(z) + b(z^N)d(z)z^{MN}, \qquad \tilde{b}(z^N)c(z) - \tilde{a}(z^N)d(z)z^{MN} \tag{10}$$

  is a GCP of length $2MN$.

  Golay's *interleaving* construction can be stated as:

$$a(z^{2N})c(z^2) + b(z^{2N})d(z^2)z, \qquad \tilde{b}(z^{2N})c(z^2) - \tilde{a}(z^{2N})d(z^2)z \tag{11}$$

  is a GCP of length $2MN$.

Repeated application of Turyn's construction, beginning with pairs of lengths 2,10 and 26 given in Section 2.4, can be used to construct GCPs for all lengths $N = 2^\alpha 10^\beta 26^\gamma$, $\alpha, \beta, \gamma \geq 0$.

2.3 Direct Constructions

In [2], Golay gave a direct construction for GCPs of length $N = 2^m$. Reference [4] gave a particularly compact description of this construction by using Algebraic Normal Forms (ANFs). With a Boolean function $a(\mathbf{x}) = \mathbf{a}(\mathbf{x_0}, \mathbf{x_1}, \ldots, \mathbf{x_{m-1}})$ in m variables, we associate a length 2^m sequence $\mathbf{a} = (a_0, a_1, \ldots, a_{2^m-1})$, where

$$a_i = (-1)^{a(i_0, i_1, \ldots, i_{m-1})}, \quad i = \sum_{k=0}^{m-1} i_k 2^k.$$

Thus the i-th term of the sequence $\mathbf{a}$ is obtained by evaluating the function a at the 2-adic decomposition of i. Then [4] showed that, for any permutation π of $\{0, 1, \ldots, m-1\}$, and any choice of constants $c_j, c, c' \in \mathbb{Z}_2$, the pair of functions

$$\begin{aligned} a(\mathbf{x}) &= \textstyle\sum_{i=0}^{m-2} x_{\pi(i)} x_{\pi(i+1)} + (\sum_{j=0}^{m-1} c_j x_j) + c \\ b(\mathbf{x}) &= a(\mathbf{x}) + x_{\pi(0)} + c' \end{aligned} \tag{12}$$

yields a length 2^m GCP $(\mathbf{a}, \mathbf{b})$.

It is simple, given this representation, to show that this construction gives a set of $m!2^m$ distinct Golay sequences of length 2^m, each of which occurs in at least 4 GCPs. Perhaps more importantly, by expressing this set in the form (12), [4] identified a large set of GCS occurring as a subset of the binary Reed-Muller code RM$(2, m)$. Consequently, each sequence in the set has PMEPR at most 2, and the Hamming distance between any two sequences in the set is at least 2^{m-2}. This set therefore has a very attractive combination of PMEPR and error-correcting properties making it applicable in OFDM applications. For further details, see [4].

It was shown in [7] that the direct construction of Golay described above and Golay's recursive constructions [2] described in Section 2.2 in fact result in the same set of Golay sequences.

2.4 Symmetry and Primitivity

We consider simple symmetry operations which leave complementary properties and sequence length of a GCP invariant.

A length N GCP $(\mathbf{a}, \mathbf{b})$ remains a GCP under the operations [2]:

- Swapping the sequences in the pair.
- Negation (element multiplication by -1) of either or both sequences.
- Reversal of either or both sequences.
- The 'linear offset' transformation $a_i \to (-1)^i a_i$, $b_i \to (-1)^i b_i$.

The action of these symmetries on a GCP generates a set of GCPs of size at most 64, which we call the *conjugates* of the orginal GCP. These symmetries can aid in computer searches for new GCPs.

We define a *primitive GCP* to be one that cannot be constructed from any shorter GCP by means of the recursive constructions described in Section 2.2. Primitive GCPs are only known to exist for lengths $2, 10$, and 26. There is one primitive pair for lengths 2 and 26, and two primitive pairs for length 10, up to equivalence via the above symmetry operations. These primitive pairs are as follows:

$$(++, \quad +-);$$
$$(-++-+-+++-, \quad -++++++--+); \qquad (+-+-++++--, \quad ++++-++--+);$$
$$(+-++--+----+-+----++---+-+, -+--++-++++-------++---+-+).$$

Golay points out that the two GCPs of length 10 are equivalent under decimation [2]. Specifically, the second pair of length 10 above is obtained from the first pair by taking successive 3rd sequence elements, cyclically. There is no proof that more primitive pairs cannot exist for $N > 100$ but none have been discovered for 40 years and it is conjectured that all GCPs arise from the four primitive pairs of lengths 2,10, and 26, as given above [1].

[1]However a very recent paper by Borwein and Ferguson [17] also regards a length 20 pair as primitive, specifically: $\{++++-+---++--++-+--+, \quad ++++-+++++---+-+-++-\}$

2.5 Enumeration

There are two main types of enumeration possible. One can enumerate the number of GCPs of a given length. Secondly, one can enumerate the number of Golay *sequences* of a given length. Since a Golay sequence is present in more than one GCP, the number of the former is greater than the number of the latter. As we've already seen in our brief discussion of OFDM, the enumeration of Golay sequences is of some practical importance. Table 1 provides a complete enumeration of GCPs for all possible lengths up to 100.

Table 1: The Number of Golay Complementary Pairs for all Lengths, $N < 100$

N	1	2	4	8	10	16	20	26	32	40	52	64	80
#GCPs[17]	4	8	32	192	128	1536	1088	64	15360	9728	512	184320	102912

From Table 1 it is evident that the largest sets occur for lengths which have a large power of 2 as a factor. Here is a useful enumeration theorem:

Theorem 1 *[2] There are exactly $2^{m+2}m!$ GCPs of length 2^m that can be derived from the primitive pair $\{++, +-\}$ by repeated application of the symmetry operations and Golay's recursive constructions.*

Next we consider the enumeration of Golay sequences. We have:

Theorem 2 *[18, 4] Golay's direct construction produces exactly $m!2^m$ Golay sequences of length 2^m.*

It was shown in [7] that the set of sequences in the above theorem is identical to that which can be obtained from the primitive pair $(++, +-)$ by repeated application of Golay's recursive constructions. Theorem 2 accounts for all Golay sequences of lengths 2^m when $1 \leq m \leq 6$. It is not known if every Golay sequence of length 2^m must arise from Golay's direct construction when $m \geq 7$.

3 The Merit Factor of Complementary Sequences

The merit factor is a useful measure of the quality of sequences in certain applications where aperiodic correlations are important. It was introduced by Golay in [18]. Let $\mathbf{a}$ be any length

N sequence. Then the merit factor of $\mathbf{a}$ is defined to be:

$$F(\mathbf{a}) = \frac{N^2}{2\sum_{k=1}^{N-1} |\rho_{\mathbf{a}}(k)|^2} \tag{13}$$

where $\rho_{\mathbf{a}}(k)$ is the AACF of $\mathbf{a}$.

The merit factor is in fact a spectral measure – it measures the mean-square deviation from the flat Fourier spectrum. Specifically,

$$1/F(\mathbf{a}) = \frac{1}{N^2}\int_0^1 (|a(e^{j2\pi t})|^2 - N)^2 dt \tag{14}$$

where $a(z)$ is a polynomial whose values on $|z| = 1$ gives the Fourier Transform of $\mathbf{a}$.

It is desirable to find sequences with high merit factor. A random sequence has merit factor around 1. It has been established that the asymptotic merit factor of a length 2^m Golay-Rudin-Shapiro (RuS) sequence is 3.0, which is high [19]. This is not the best possible, for instance, shifted-Legendre sequences attain an asymptotic merit factor of 6.0 [20], and computer searches up to length 200 have revealed ± 1-sequences of merit factor around 8.5. There is also the celebrated Barker sequence of length 13 which has merit factor 14.08. However the length $N = 2^m$ RuS sequences $\mathbf{a}_m$ are notable as the quantities $\sigma_m = \sum_{k=0}^{2^m-1} |\rho_{\mathbf{a}_m}(k)|^2$ obey a simple Generalised Fibonacci Recursion, namely,

$$\sigma_m = 2\sigma_{m-1} + 8\sigma_{m-2} \tag{15}$$

with initial conditions $\sigma_1 = 1$, $\sigma_2 = 2$ [19].

This recursion immediately gives an asymptotic merit factor for the RuS sequences of 3.0 and is significant because it demonstrates the existence of a sequence family with large merit factor for which the merit factors do not need to be computed explicitly. This asymptotic value of 3.0 also holds for any Golay sequence obtained by applying symmetry operations to the RuS sequences [19]. Taking any other non-Golay pair as a starting seed always gives an asymptotic merit factor of K, for some constant, $K < 3.0$ [21].

4 Low-Complexity Correlation

The pairwise property of GCPs has been exploited for channel measurement [1] but, until 1990 or thereabouts the properties of individual sequences of a GCP were not wholly exploited [22], although Shapiro had stated equation (4) in his master's thesis of 1951 [9]. Budisin argues that Golay sequences are as good, if not better, than m-sequences for application as pseudo-noise sequences, due to their superior aperiodic spectral properties [22]. Also, there are more Golay sequences than m-sequences. Fig 1 lends some support for this

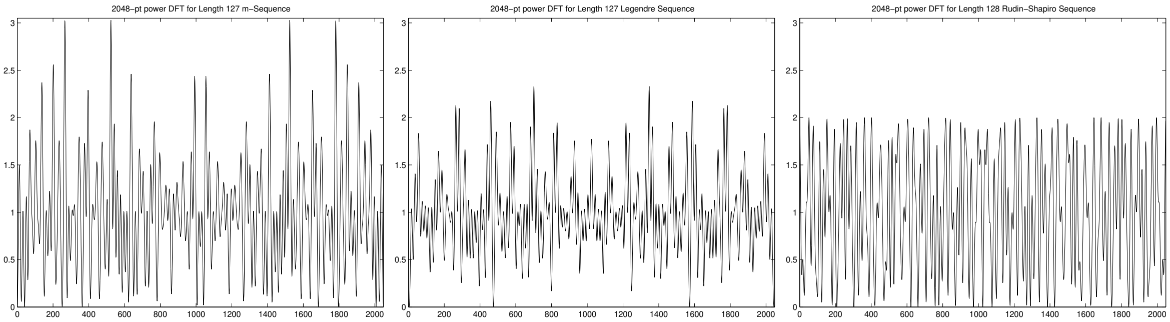

Figure 1: Power Spectra for Length 127 m-Sequence, Length 127 Shifted-Legendre, and Length 128 Rudin-Shapiro Sequences, (Power on y-axis, Spectral Index on x-axis)

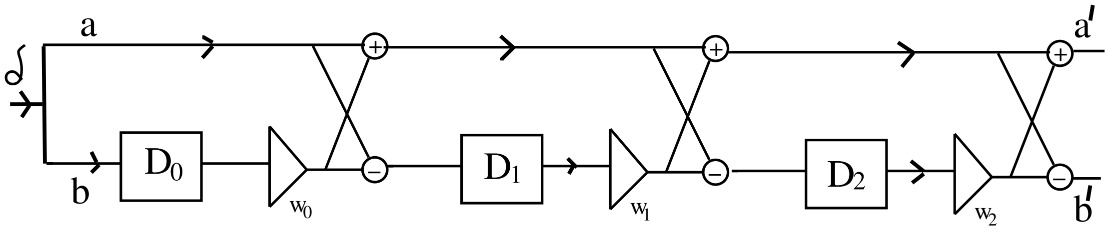

Figure 2: Fast Golay Correlator

view, where the Fourier spectra (from left to right) of a length 127 m-sequence and a length 127 shifted-Legendre sequence are compared with that of a length 128 RuS sequence.

Budisin [22] proposed a highly efficient method to perform correlation of an incoming data stream with a Golay sequence of length N which achieves a complexity of $2\log_2(N)$ operations per sample, as opposed to N operations per sample for direct correlation. To do this he interpreted the Golay construction for length $N = 2^m$ sequences using delay to implement concatenation, i.e. $\mathbf{a}|\mathbf{b}$ can be implemented as $\mathbf{a}[k]+\mathbf{b}[k+D]$, where $[k]$ indicates the starting time of $\mathbf{a}$ and D is the length (duration) of $\mathbf{a}$. Implementing the recursion of (8) is then achieved by serially combining delay elements, D_i, of duration 2^i, as shown in Fig 2. Now we commence our Rudin-Shapiro recursion with $\mathbf{a} = \mathbf{b} = 1$. In other words, we input the δ function to the left-hand side of Fig 2, and output our GCP, $(\mathbf{a}', \mathbf{b}')$, on the right. So the pair of Golay sequences realised by Fig 2 are two impulse responses. We can therefore re-interpret Fig 2 as a filter which correlates a received sequence, input from the left, with the reversals of $\mathbf{a}$ and $\mathbf{b}$. By choosing the ω_i in Fig 2 from $\{1, -1\}$, we can choose to correlate with different length 2^m GCPs.

5 Complementary Sets and Orthogonal Matrices

5.1 Complementarity with respect to a set of sequences

Golay complementary pairs can be generalised to sets containing two or more sequences [23]. Analogously to (2), we say that a set of T bi-polar sequences of length N $(\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_T)$ form a Complementary Sequence Set of size T (a T-CSS) if

$$\sum_{i=1}^{T} \rho_{\mathbf{a}_i}(k) = 0 \qquad k \neq 0. \tag{16}$$

In terms of polynomials, this is equivalent to

$$\sum_{i=1}^{T} |a_i(z)|^2 = TN, \qquad |z| = 1. \tag{17}$$

Thus all the DFT components of a sequence $\mathbf{a}$ that lies in a T-CSS are of size at most $\sqrt{TN}$. Of course, a 2-CSS is just a Golay complementary pair.

To date, little work has been done to formally establish primitivity conditions for T-CSS, $T > 2$, although Golay already found 4-CSS in [1]. It can be shown that CSS only occur for T even, and Turyn showed that 4-CSS are only admissable at lengths N if N is a sum of at most *three* squares [14]. Dokovic later showed that 4-CSS exist for all even $N < 66$ [24]. [23] showed that for a CSS of odd length N, T must be a multiple of 4. By way of example, here are 4-CSS of lengths 3,5,7:

$$\{+++, \quad -++, \quad +-+, \quad ++-\}$$
$$\{+----, \quad -++-+, \quad +---+, \quad ---+-\}$$
$$\{+++-+++, \quad +-+++--, \quad +--+-++, \quad ++-+---\}$$

Thus 4-CSS can exist at lengths N where 2-CSS cannot. Turyn presented constructions for 4-CSS for all odd lengths $N \leq 33$, and $N = 59$ [14].

An orthogonal matrix is defined to be a matrix whose columns are pairwise orthogonal. The following Theorem is straightforward.

Theorem 3 *[23] Let $\mathbf{P}$ be a $T \times N$ orthogonal matrix, $N \leq T$. Then the rows of P form a T-CSS of length N.*

The primitive GCP $(++, +-)$, is an example of Theorem 3. When the elements of $\mathbf{P}$ are ± 1 and $N = T$, then $\mathbf{P}$ is a Hadamard matrix, so a subset of T-CSS is provided by the set of Hadamard matrices.

5.2 Symmetries and Constructions

The symmetries and constructions for GCPs given in Sections 2.2 and 2.3 generalise to give constructions for T-CSS. One can also construct T'-CSS by combining T-CSS, $T' > T$. As one example, we have:

- Let $(\mathbf{u}, \mathbf{v})$ and $(\mathbf{x}, \mathbf{y})$ be GCPs of length N_0 and N_1, respectively. Then, $(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d})$ is a 4-CSS of length $N_0 + N_1$, where

$$\mathbf{a} = (\mathbf{u}|\mathbf{x}), \quad \mathbf{b} = (\mathbf{u}|-\mathbf{x}), \quad \mathbf{c} = (\mathbf{v}|\mathbf{y}), \quad \mathbf{d} = (\mathbf{v}|-\mathbf{y})$$

A fundamental recursive construction generalising Golay's recursive constructions and relating CSS to orthogonal matrices is given in [23]:

Theorem 4 *[23] Let $(\mathbf{a_0}, \mathbf{a_1}, \ldots, \mathbf{a_{T-1}})$ be a T-CSS of length N, represented by a $T \times N$ matrix, $\{\mathbf{F}\}$, with rows $\mathbf{a_j}$. Let $\mathbf{O} = (o_{ik})$ be an $S \times T$ orthogonal matrix (so $S \geq T$). Define*

$$\mathbf{F}' = \mathbf{F} \odot \mathbf{O} = \begin{pmatrix} o_{00}\mathbf{a_0} & o_{01}\mathbf{a_1} & \ldots & o_{0(T-1)}\mathbf{a_{T-1}} \\ o_{10}\mathbf{a_0} & o_{11}\mathbf{a_1} & \ldots & o_{1(T-1)}\mathbf{a_{T-1}} \\ \ldots & \ldots & \ldots & \ldots \\ o_{(S-1)0}\mathbf{a_0} & o_{(S-1)1}\mathbf{a_1} & \ldots & o_{(S-1)(T-1)}\mathbf{a_{T-1}} \end{pmatrix} \tag{18}$$

Then $\mathbf{F}'$ is an $S \times TN$ matrix whose rows form an S-CCS of length TN.

Taking $(\mathbf{a}_0, \mathbf{a}_1)$ to be a GCP and $O = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, we recover Golay's concatenation construction. The basic symmetry operations can be interpreted as row/column permutations of $\mathbf{O}$ and as point-multiplication of rows of $\mathbf{O}$ by a constant vector (these operations maintain the orthogonality of $\mathbf{O}$).

The following theorem combines T-CSS of different lengths to build $T' - CSS$, where $T' > T$.

Theorem 5 *[25] Suppose there exist T_0-CSS, T_1-CSS,. . .,T_{t-1}-CSS, of lengths $N_0, N_1, \ldots, N_{t-1}$, respectively. Let $T =$ lcm$(T_0, T_1, \ldots, T_{t-1})$. Suppose there also exists an $S \times T$ orthogonal matrix with ± 1 entries. Then there exists an ST-CSS of length $N' = N_0 + N_1 + \ldots + N_{t-1}$.*

As with GCPs, CSS also have applications to OFDM: a primary drawback with the proposal to use the set of length 2^m GCPs as a codeset for OFDM is that the code rate of the set rapidly decreases as m increases. To obtain a larger codeset one can consider sequences that lie in T-CSS for some $T > 2$. The resulting codeset will have PMEPR at most T.

5.3 Complementarity With Respect to a Larger Set of Transforms

Although CS pairs and, more generally, CSS, are usually defined to be complementary with respect to their AACFs (with a corresponding property on power spectra under the one-dimensional DFT), one can, more generally, define and discover sets which are complementary with respect to *any* specified transform. It can be shown, [26], that Golay CSS of length 2^m, as constructed using Theorem 5, have a very strong property:

Theorem 6 *Let* $\mathbf{U}$ *be a* 2×2 *complex-valued matrix such that* $\mathbf{U}\mathbf{U}^\dagger = 2\mathbf{I}$*, where* $\mathbf{I}$ *is the* 2×2 *identity matrix, '†' means transpose-conjugate, and the elements of* $\mathbf{U}$*,* u_{ij}*, satisfy* $|u_{00}| = |u_{01}| = |u_{10}| = |u_{11}|$*. Let* $\{\mathbf{U}_k, 0 \leq k < m\}$ *be a set of any* m *of these matrices. Define* $\mathbf{M} = \mathbf{U}_0 \otimes \mathbf{U}_1 \otimes \cdots \otimes \mathbf{U}_{m-1}$*. Let* $N = 2^m$ *and let* $(\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_{T-1})$ *be any* T*-CSS of length* N *constructed by Theorem 5. Finally, let* $\mathbf{A}_i^M = \mathbf{M}\mathbf{a}_\mathbf{i}$ *be the* N*-point spectrum of* $\mathbf{a}$ *with respect to* $\mathbf{M}\mathbf{a}$ *with elements,* $A_{k,i}^M$*,* $0 \leq k < N$*. Then:*

$$\sum_{i=0}^{T-1} |A_{k,i}^M|^2 = TN. \tag{19}$$

Theorem 6 implies that

$$|A_{k,i}^M|^2 \leq TN \qquad \forall k, i, T, \mathbf{M} \tag{20}$$

The combined set of all rows of all possible transform matrices, $\mathbf{M}$, includes the one-dimensional DFT and the Walsh-Hadamard Transform (WHT), along with infinitely many other transforms.

As just one example of the application of this result, recall that cryptographers typically perform linear cryptanalysis of cipher components by looking for peaks in the WHT spectrum, see for example [27]. It is known that for m even, length 2^m GCPs are bent, that is, have a completely flat WHT spectrum [28]. [26] shows that Golay constructions generate a large set of GCPs and CSS which also have a relatively flat spectrum with respect to the WHT, amongst other transforms. So Golay CSS may have applications to cryptography.

5.4 CSS Mates

Let $\mathbf{A} = (\mathbf{a}_0, \mathbf{a}_1, \ldots, \mathbf{a}_{T-1})$ and $\mathbf{B} = (\mathbf{b}_0, \mathbf{b}_1, \ldots, \mathbf{b}_{T-1})$ be two T-CSS. Then $\mathbf{A}$ and $\mathbf{B}$ are called 'mates' if $\mathbf{a}_i$ and $\mathbf{b}_i$ are orthogonal as vectors for each i. We say that $\mathbf{A}$ and $\mathbf{B}$ are 'mutually orthogonal CSS' (although, in general, $\mathbf{a}_i$ is not orthogonal to $\mathbf{b}_j$, $i \neq j$). Sets $(\mathbf{A}_0, \mathbf{A}_1, \ldots, \mathbf{A}_{U-1})$ of pairwise mutually orthogonal CSS can be recursively constructed in a similar way to CSS [23].

6 Complementary Sequences over Larger Alphabets

Virtually all the symmetries and constructions mentioned so far for bi-polar sequences can be generalised to sequences over other alphabets, but note that autocorrelation is now modified to include conjugacy, i.e. (1) is modified to,

$$\rho_{\mathbf{a}}(k) = \sum_{i=0}^{N-k-1} a_i a_{i+k}^*, \qquad 0 \leq k \leq N-1. \tag{21}$$

where * means 'complex conjugate'.

For quadriphase CSS, the symmetry operations generate an equivalence class of up to 1024 sequences [29]. For polyphase pairs, unlike the GCP case, there is no restriction that the length must be the sum of two squares. Sivaswamy and Frank investigated and discovered many polyphase CSS, including those of odd length [30, 31]. The simplest polyphase T-CSS of length T is formed from the rows of the T-point DFT matrix. In fact, from Theorem 3, the rows of any $T \times T$ orthogonal polyphase matrix form a T-CSS of length T. [30] identified a length 3 quadriphase primitive pair, $(002,\ \ 010)$ (where $0, 1, 2, 3$ mean i^0, i^1, i^2, i^3, respectively), derived quadriphase versions of GCPs and synthesized sequence pairs of lengths 3.2^k. [31] further presented the following primitive quadriphase Golay pairs, of lengths 5 and 13: $(01321, 00013)$, and $(0001200302031, 0122212003203)$. Note that the lengths here, 5 and 13, are half the length of the lengths 10 and 26 primitive bi-polar GCPs, but no transform is known between the sets.

[4] and [7] constructed many CSS with phase alphabet 2^h and any even phase alphabet, respectively. To do so, they worked with non-binary generalisations of the Reed-Muller codes. The resulting sequences have application to OFDM with non-binary modulation.

Many polyphase 3-CSS exist. For instance, [31] presented the triphase 3-CSS (01110,11210,00201), and provided a (possibly non-exhaustive) list of lengths, N, for which a CSS exists, N up to 100:

Polyphase 3-CSS:

1-22,24-27,30,32,33,36,37,39-42,45,48,49,51-54,57,58,60,61,63-66,72,73,75,78,80,81,90,96,97,100

Polyphase 4-CSS: All lengths except 71,89

$T > 4$: All lengths.

A recent exhaustive search [29] found *all* quadriphase Golay pairs up to length $N = 13$. These are summarised below where only those for which no construction is known have been counted. These are the *possible* primitive pairs. The figures also omit the GCPs which are a subset of quadriphase pairs:

Length	2	3	4	5	6	7	8	9	10	11	12	13
#Inequivalent Pairs[29]	–	1	–	1	2	–	4	–	14	1	32	1

Golay pairs over the alphabet $\{0, 1, -1\}$ have been found for all lengths, N. For such a set, the weight of the set becomes an important extra parameter. The weight W is the sum of the in-phase AACF coefficients, i.e. for a set $(\mathbf{a}_j)$, we have $W = \sum_j \rho_0(\mathbf{a}_j)$. For instance, here is a 4-CSS of weight 7 and length 7 over the alphabet $\{0, 1, -1\}$:

$$\{+000+00, \quad 0+0+0-0, \quad 00+0000, \quad 000000+\}.$$

The larger W, the closer is the CSS to one over a bi-polar alphabet.

Yet more CSS can be found by considering multilevel and QAM alphabets.

7 Hadamard Matrices from Complementary Sequences

In Section 5 we showed that Hadamard matrices can be used to construct CSS. The converse is also true: CSS can be used to construct Hadamard matrices. Here, we present the two most well-known constructions, where Theorems 7 and 8 use the *periodic* complementary property of a complementary set (see Section 8).

Theorem 7 *[32] Let $(\mathbf{a}, \mathbf{b})$ be a GCP of length N. Let $\mathbf{A}$ and $\mathbf{B}$ be $N \times N$ circulant matrices with first rows $\mathbf{a}$ and $\mathbf{b}$, respectively. Then,*

$$\begin{pmatrix} \mathbf{A} & -\mathbf{B} \\ \mathbf{B^T} & \mathbf{A^T} \end{pmatrix} \text{ is a } 2N \times 2N \text{ Hadamard matrix}$$

Theorem 7 can be generalised to quadriphase Hadamard matrices by making $(\mathbf{a}, \mathbf{b})$ a quadriphase Golay pair, and by substituting conjugation for transpose.

Theorem 8 *[14, 15] Let $(\mathbf{u}, \mathbf{v})$ and $(\mathbf{x}, \mathbf{y})$ be GCPs of lengths N_0 and N_1, respectively. Then $\mathbf{a} = \mathbf{u}|\mathbf{x}$, $\mathbf{b} = \mathbf{u}|-\mathbf{x}$, $\mathbf{c} = \mathbf{v}|\mathbf{y}$, and $\mathbf{d} = \mathbf{v}|-\mathbf{y}$ form a length $N = N_0 + N_1$ 4-CSS. Let $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ be $N \times N$ circulant matrices with first rows $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$, respectively. Let $\mathbf{R}$ be a back-circulant $N \times N$ permutation matrix. Then,*

$$\begin{pmatrix} \mathbf{A} & -\mathbf{BR} & -\mathbf{CR} & -\mathbf{DR} \\ \mathbf{BR} & \mathbf{A} & -\mathbf{D^TR} & \mathbf{C^TR} \\ \mathbf{CR} & \mathbf{D^TR} & \mathbf{A} & -\mathbf{B^TR} \\ \mathbf{DR} & -\mathbf{C^TR} & \mathbf{B^TR} & \mathbf{A} \end{pmatrix} \text{ is a } 4N \times 4N \text{ Goethals-Seidel (Hadamard) matrix.}$$

8 Periodic and Odd-Periodic (Negaperiodic) Complementary Sequences

Researchers have recently become interested in constructing *periodic* and/or *odd-periodic* CSS. Periodic CSS were considered in [33].

Using polynomial form, Periodic Autocorrelation (PACF) of the length N sequence $\mathbf{a}$ can be expressed as,

$$\text{PACF}(a(x)) = \left\langle a(x)a(x^{-1})\right\rangle_{x^N-1}$$

where '$\langle * \rangle_M$' reduces $*$ mod M.

Similarly, Negaperiodic (odd-periodic) Autocorrelation (NACF) can be expressed as,

$$\text{NACF}(a(x)) = \left\langle a(x)a(x^{-1})\right\rangle_{x^N+1}$$

There is a simple relationship relating aperiodic, periodic, and odd-periodic AACF, as follows:

Let $a(x)$ represent a length N sequence. Then the AACF of $\mathbf{a}$ can be computed, via the Chinese Remainder Theorem, as,

$$a(x)a(x^{-1}) = \frac{x^N+1}{2}\left\langle a(x)a(x^{-1})\right\rangle_{x^N-1} - \frac{x^N-1}{2}\left\langle a(x)a(x^{-1})\right\rangle_{x^N+1} \tag{22}$$

(22) expresses AACF in terms of PACF and NACF. It follows that,

- A T-CSS is also a periodic and a negaperiodic T-CSS.
- A set of length T sequences is only a T-CSS if it is both a periodic and negaperiodic T-CSS.

Periodic and negaperiodic CSS can be used instead of, say, m-sequences, for their desirable correlation and spectral properties. They are much easier to find than aperiodic CSS due to their algebraic structure via embedding in a finite polynomial ring. Moreover, in a search for aperiodic CSS, an initial sieve can be undertaken by first identifying periodic and negaperiodic CSS, and then looking for the intersection of the two sets.

It is known that periodic GCP do not exist for lengths 36 and 18, respectively.

We have the following theorem,

Theorem 9 *[34] If a length N periodic GCP exists, such that $N = p^{2l}u$, $p \neq u$, p prime, $p = 3(mod 4)$, then $u \geq 2p^l$.*

Dokovic discovered a periodic GCP of length 34 [24]. This is significant because no aperiodic GCP exists at that length. The next unresolved case for periodic GCPs is at length 50. [33] and [24] also present many T-CSS for $T > 2$.

Lüke [35] has found many odd-periodic GCPs for even lengths N where an aperiodic GCP cannot exist, e.g. $N = \frac{p^u+1}{2}$, p an odd prime, and also for all even N, $N < 50$. But he did not find any odd-length pairs.

References

[1] M.J.E. Golay, "Multislit spectroscopy," *J. Opt. Soc. Amer.*, **39**, pp. 437–444, 1949.

[2] M.J.E. Golay, "Complementary series," *IRE Trans. Inform. Theory*, **IT-7**, pp. 82–87, Apr. 1961.

[3] M. Nazarathy, S.A. Newton, R.P. Giffard, D.S. Moberly, F. Sischka and W.R. Trutna, Jr., "Real-time long range complementary correlation optical time domain reflectometer," *IEEE J. Lightwave Technology*, **7**, no. 1, pp. 24–38, 1989.

[4] J.A. Davis, J. Jedwab, "Peak-to-mean power control in OFDM, Golay complementary sequences and Reed-Muller codes," *IEEE Trans. Inform. Theory*, **IT-45**, no. 7, pp. 2397–2417, Nov. 1999.

[5] R.D.J. van Nee, "OFDM codes for peak-to-average power reduction and error correction," in *IEEE Globecomm 1996* (London, U.K., Nov. 1996), pp. 740–744.

[6] H. Ochiai and H. Imai, "Block coding scheme based on complementary sequences for multicarrier signals," *IEICE Trans. Fundamentals*, **E80-A**, pp. 2136–2143, 1997.

[7] K.G. Paterson, "Generalized Reed-Muller codes and power control in OFDM modulation," *IEEE Trans. Inform. Theory*, **IT-46**, no. 1, pp. 104–120, Jan. 2000.

[8] P. Spasojevic,C.N. Georghiades, "Complementary sequences for ISI channel estimation," *IEEE Trans. Inform. Theory*, **IT-47**, pp. 1145–1152, March 2001.

[9] H.S. Shapiro, "Extremal Problems for Polynomials," *M.S. Thesis, M.I.T.*, 1951.

[10] J. Beck, "Flat polynomials on the unit circle – note on a problem of Littlewood,", *Bull. London Math. Soc.*, **23**, pp. 269–277, 1991.

[11] J. Kahane, "Sur les polynomes à coefficients unimodulaires," *Bull. London Math. Soc.*, **12**, pp. 321–342, 1980.

[12] J.E. Littlewood, "On polynomials $\sum \pm z^m$, $\sum \exp(\alpha_m) z^m$,$z = e^{i\theta}$," *J. London Math. Soc.*, **41**, pp. 367–376, 1966.

[13] W. Rudin, "Some theorems on Fourier coefficients," *Proc. Amer. Math. Soc.*, **10**, pp. 855–859, 1959.

[14] R. Turyn, "Hadamard matrices, Baumert-Hall units, four-symbol sequences, pulse compression, and surface wave encodings," *J. Comb. Theory Ser. A*, **16**, pp. 313–333, 1974.

[15] C.H. Yang, "Hadamard matrices, finite sequences, and polynomials defined on the unit circle," *Mathematics of Computation*, **33**, 146, pp. 688–693, April 1979.

[16] S. Eliahou, M. Kervaire, B. Saffari, "A new restriction on the lengths of Golay complementary sequences," *Journ. Comb. Theory Ser. A* , **55**, pp. 49–59, 1990.

[17] P. Borwein,R.A. Ferguson, (2000) "A complete description of Golay pairs for lengths up to 100," *in preparation*, preprint available [Online]. Simon Fraser University, http://www.cecm.sfu.ca/∼pborwein/, [May, 2002]

[18] M.J.E. Golay, "Seives for low autocorrelation binary sequences," *IEEE Trans. Inform. Theory*, **IT-23**, no. 1, pp. 43–51, 1977.

[19] T. Høholdt, H.E. Jensen and J. Justesen, "Aperiodic correlations and merit factor of a class of binary sequences," *IEEE Trans. Inform. Theory*, **IT-31**, pp. 549–552, Jul. 1985.

[20] T. Høholdt and H.E. Jensen, "Determination of the merit factor of Legendre sequences," *IEEE Trans. Inform. Theory*, **IT-34**, pp. 161–164, Jan. 1988.

[21] P. Borwein,M. Mossinghoff, "Rudin-Shapiro like polynomials in L_4", *Math. Comp.*, **69**, pp. 1157–1166, 2000.

[22] S.Z. Budisin, "Efficient pulse compressor for Golay complementary sequences," *Elec. Lett.*, **27**, no. 3, pp. 219–220, 31st Jan. 1991.

[23] C.-C. Tseng, C.L. Liu, "Complementary sets of sequences," *IEEE Trans. Inform. Theory*, **IT-18**, no. 5, pp. 644–651, Sept. 1972.

[24] D.Z. Dokovic, "Note on periodic complementary sets of binary sequences," *Designs, Codes and Cryptography*, **13**, pp. 251–256, 1998.

[25] K. Feng,P.J-S. Shiue,Q. Xiang, “On aperiodic and periodic complementary binary sequences”, *IEEE Trans. Inf. Theory*, **45**, 1, pp. 296–303, Jan 1999.

[26] M.G. Parker, C. Tellambura, “A construction for binary sequence sets with low peak-to-average power ratio,” *Int. Symp. Inform. Theory, Lausanne, Switzerland*, June 30-July 5, 2002.

[27] M. Matsui, “Linear cryptanalysis method for DES,” in *Advances in Cryptology – Eurocrypt93*, Lecture Notes in Computer Science vol. 765, pp. 386–397, Springer, 1993.

[28] F.J. MacWilliams and N.J.A. Sloane, *The theory of error-correcting codes*, Amsterdam, North-Holland, 1977.

[29] W.H. Holzmann,H. Kharaghani, “A computer search for complex Golay sequences,” *Aust. Journ. Comb.*, **10**, pp. 251-258, 1994.

[30] R. Sivaswamy, “Multiphase complementary codes,” *IEEE Trans. Inform. Theory*, **IT-24**, pp. 546–552, 1978.

[31] R.L. Frank, “Polyphase complementary codes,” *IEEE Trans. Inform. Theory*, **IT-26**, no. 6, pp. 641–647, Nov. 1980.

[32] C.H. Yang, “On Hadamard matrices constructible by circulant submatrices,” *Math. Comp.*, **25**, pp. 181–186, 1971.

[33] L. Bomer, M. Antweiler, “Periodic complementary binary sequences,” *IEEE Trans. Inform. Theory*, **IT-36**, pp. 1487–1494, 1990.

[34] K.T. Arasu,Q. Xiang, “On the existence of periodic complementary binary sequences,” *Designs, Codes and Cryptography*, **2**, 3, pp. 257–262, 1992.

[35] H.D. Lüke, “Binary odd-periodic complementary sequences,” *IEEE Trans. Inform. Theory*, **IT-43**, no. 1, pp. 365–367, Jan. 1997.