A. Milewski

Periodic Sequences with Optimal Properties for Channel Estimation and Fast Start-Up Equalization

The problems of fast channel estimation and fast start-up equalization in synchronous digital communication systems are considered from the viewpoint of the optimization of the training sequence to be transmitted. Various types of periodic sequences having uniform discrete power spectra are studied. Some of them are new and may be generated with data sets commonly used in phase modulation systems. As a consequence of their power spectra being flat, these sequences ensure maximum protection against noise when initial equalizer settings are computed via channel estimates and noniterative techniques.

Introduction

Sequences (or "codes") with good autocorrelation properties have been studied in communications literature for over twenty years because of their applications to radar and the synchronization of communications systems [1, 2]. For a current and more general reference on this subject, see Alltop [3]. More recently, "training sequences" with similar properties have been studied and used for fast start-up equalization [4–7].

Telephone lines present large amounts of linear amplitude and phase distortion. Fast turnaround is an important element of modem performance since messages (especially on multidrop lines) are often short and resynchronization is required frequently. The use of training sequences permits rapid equalization of the transmission channel without any prior knowledge of signal distortions, provided that they are not too extreme (e.g., they present no spectral nulls). Performing equalization every time resynchronization is required has the advantage of simplifying the overall control and maintenance procedures. For example, in the case of multidrop lines, it avoids the necessity of stocking the equalizer coefficients corresponding to each secondary modem connected to the line, and hence it avoids the procedures for setting and maintaining these coefficients in case of line changes or variations.

All training sequences share two properties:

1. Their autocorrelation function is small, except at the origin.
2. Some limitation is imposed on the numerical values taken by the sequences.

The first property is required to make these sequences as nearly as possible "impulse-equivalent." The second is due to the requirement for peak amplitude limitation in all practical implementations.

The main distinction among the many types of sequences is whether they are periodic or not, since this affects the definition of the autocorrelation function and the spectral properties. In this paper we confine ourselves to the subject of periodic sequences, because of their particular suitability for fast start-up equalization. More precisely, we discuss complex sequences having *constant amplitude and zero autocorrelation* (CAZAC sequences) and compare their performance with the better known "maximal-length" sequences.

The channel model

Let us consider a complex channel with additive white noise and finite impulse response which is sampled at T intervals. The channel input at time nT is a complex number u_n, and the channel output y_n is

$$y_n = \sum_{k=N_1}^{N_2} r_k u_{n-k} + w_n , \qquad (1)$$

where $\mathbf{r}$ is a complex vector of length $L = N_2 - N_1 + 1$, defining the channel impulse response, and w_n is a complex random variable with the expected values

$$E(w_n \bar{w}_m) = \sigma^2 \quad \text{if} \quad n = m,$$
$$= 0 \quad \text{if} \quad n \neq m, \text{ and}$$
$$E(w_n) = 0 \quad \text{for all } n.$$

This channel model can be considered as an idealization of a quadrature-amplitude-modulated (QAM) sampled channel.

The spectrum of this idealized channel is given by

$$\sum_{k=N_1}^{N_2} r_k e^{2\pi j f k T},$$

a Fourier series of L coefficients, where f is the frequency and T is the sampling interval. Since the spectrum of a sampled channel is a periodic function of f, it can be approximated as closely as is needed by taking L large enough.

Training sequences for fast equalization

Fast start-up digital equalization requires rapid, accurate, and dependable estimation of channel characteristics (represented by r_n), and also rapid and hardware-implementable calculation of the equalizer coefficients from the data furnished by the estimation.

There are two main advantages to using a training sequence for the purpose of estimating the channel characteristics. The transmitted symbol is known to the receiver (thus detection errors are eliminated), and the training sequence can be chosen to have certain desirable properties. Among these are that the estimation method is hardware-implementable and that the process is insensitive to noise.

An important property of a training sequence is the length of its period. In general, the longer period leads to the better channel estimate. However, it is important to note that if there is no noise and if the channel response is of finite length, then it can be completely estimated using a training sequence period equal to this length.

A well-known example of such a training sequence is the sending of a single unit pulse every L-baud interval. The r_n are then estimated directly from the received signal. The disadvantage of this training sequence is its low power and its resultant sensitivity to noise.

• *Calculating the channel response in the absence of noise*
In this section we assume that $w_n = 0$ and that the transmitted symbols u_n come from a training sequence of period L, such that L is equal to the length of the channel response; that is, $r_n = 0$ if $n < N_1$ or $n > N_2$, and $L = N_2 - N_1 + 1$.

Then the ith received signal can be written

$$x_i = \sum_{n=N_1}^{N_2} u_{i-n} r_n . \tag{2}$$

If L successive signals are observed, we shall obtain L linear equations in L unknowns and so the r_n are in principle known if and only if the matrix $\mathbf{M} = (u_{i-j})$ is nonsingular. Note that the training sequence of the preceding subsection corresponds to $\mathbf{M} = \mathbf{I}_L$ (the unit matrix of order L). Note also that if the training sequence is started at the instant $i = 0$, then the first x_n satisfying (2) is x_{N_2}. Since we require L observations, complete knowledge of the channel can be obtained $(N_2 + L - 1)$-baud intervals after the start of the training sequence, provided that it is continued during N_1-baud intervals.

We now consider the problem of finding the r_n from the set of equations (2). The periodicity of the training sequence permits a computationally efficient solution, since the matrix $\mathbf{M}$ is a circulant matrix and matrix multiplication can be replaced by periodic convolution. Thus (2) can be rewritten

$$\mathbf{x} = \mathbf{u} * \mathbf{r}, \tag{3}$$

where $\mathbf{x}$ is the vector consisting of L samples of the received signal, etc.

Denoting the discrete Fourier transform (DFT) of $\mathbf{x}$ by X, etc., we have $X = U \times R$ and so $\mathbf{r} = \mathrm{IDFT}\,(X/U) = \mathrm{IDFT}(V \times X)$ or

$$\mathbf{r} = \mathbf{v} * \mathbf{x}, \tag{4}$$

where $\mathbf{v}$ is the IDFT (inverse discrete Fourier transform) of $V = 1/U$. Since the components of U are the eigenvalues of $\mathbf{M}$, V exists if and only if $\mathbf{M}$ is nonsingular.

• *Estimating the channel response in the presence of noise*
If in (4) we replace $\mathbf{x}$ with $\mathbf{y}$, where $\mathbf{y} = \mathbf{x} + \mathbf{w}$ and $\mathbf{w}$ is a vector of noise samples, then

$$\hat{\mathbf{r}} = \mathbf{v} * \mathbf{y} = \mathbf{v} * \mathbf{x} + \mathbf{v} * \mathbf{w}, \tag{5}$$

where $\hat{\mathbf{r}}$ is the estimate of $\mathbf{r}$ given by the linear estimator $\mathbf{v}$.

Because of the assumptions about the noise, we deduce from (5) that $E(\hat{\mathbf{r}}) = \mathbf{r}$, i.e., that the estimator $\mathbf{v}$ is unbiased and that the mean square error (or error variance) is

$$E\left(\sum_i |\hat{r}_i - r_i|^2\right) = L\sigma^2 \sum_i |v_i|^2. \tag{6}$$

We now derive the condition for minimizing (6) over all training sequences of average unit power, i.e. all $\mathbf{u}$, such that

$$\sum_i |u_i|^2 = L.$$

We have, by the DFT equivalent of Parseval's formula,

$$\sum_i |v_i|^2 = \frac{1}{L} \sum_i |V_i|^2,$$

and since $V_i = 1/U_i$, we have

$$\sum_i |v_i|^2 = \frac{1}{L}\sum_i \frac{1}{|U_i|^2}; \tag{7}$$

hence the error variance resulting from the use of the sequence **u** is

$$\sigma^2 \sum_i \frac{1}{|U_i|^2}.$$

Also,

$$\sum_i |U_i|^2 = L \sum_i |u_i|^2 = L^2,$$

and it is easy to show that (7) is minimized when all the $|U_i|^2$ are equal,

$$|U_i|^2 = L \qquad i = 0, 1, \cdots, L-1,$$

that is, when the sequence **u** has zero autocorrelation. The error variance (7) then takes the value σ^2.

It is interesting to compare this result with the error variance obtained when a maximal-length training sequence taking the values +1 and −1 is employed. Such a sequence is characterized in the frequency domain by a DFT which has

$$U_0 = 1,$$

$$|U_1|^2 = |U_2|^2 = \cdots = |U_{L-1}|^2 = L + 1.$$

Its error variance therefore is

$$\sigma^2 \sum_i \frac{1}{|U_i|^2} = \sigma^2\left(1 + \frac{L-1}{L+1}\right) = \sigma^2 \frac{2L}{L+1}.$$

We see that, for large L, a maximal-length sequence is at a 3-dB disadvantage compared to a sequence with zero autocorrelation.

- *Effect of the training sequence on fast equalization*

We now consider the effect of the training sequence on calculating the equalizer coefficients. Mueller and Spaulding [5] have shown that in the *no-noise* case, perfect equalization at a discrete number of frequency points can be achieved by the coefficients

$$\mathbf{c}_0 = \mathrm{IDFT}\,\frac{U}{X} = \mathrm{IDFT}\,\frac{1}{R}. \tag{8}$$

The same method can be applied to calculating the coefficients in the presence of noise, provided that a suitable training sequence is used. For if we calculate the coefficients of the equalizer by the formula

$$\mathbf{c} = \mathrm{IDFT}\,\frac{U}{Y}, \tag{9}$$

we must ensure that the denominator

$$Y = U \times R + W \tag{10}$$

does not contain any zero terms.

Assuming that we have no information concerning R (other than that it is nonzero), it is clear that the safest training sequence is such that the U_i have constant amplitude, or in other words that **u** has zero autocorrelation.

We now evaluate the error due to periodic random data, i.e., a random vector of length L and average power = 1 repeatedly transmitted. This approximates the transmission of random data similarly to the way in which $\mathbf{c}_0$ approximates the "true" optimal coefficients $\mathbf{c}_{\mathrm{opt}}$, obtained with an infinite random sequence.

We shall use the notation

$$|\mathbf{c}|^2 \triangleq \sum_{i=0}^{L-1} |c_i|^2,$$

and denote the data vector by **d**. We first observe that the error variance due to $\mathbf{c}_0$ is

$$E_0 \triangleq |\mathbf{c}_0|^2 \times \sigma^2 = \frac{1}{L}\sum_{i=0}^{L-1} \frac{\sigma^2}{|R_i|^2}$$

Writing $\mathbf{c} = \mathbf{c}_0 + \triangle\mathbf{c}$ we have the error variance due to **c**

$$= |\mathbf{c}|^2 \times \sigma^2 + \frac{1}{L}\mathrm{E}(|\triangle\mathbf{c} * \mathbf{r} * \mathbf{d}|^2),$$

$= |\mathbf{c}|^2 \times \sigma^2 + |\triangle\mathbf{c} * \mathbf{r}|^2$, since **d** is assumed random, and

$= |\mathbf{c}_0 + \triangle\mathbf{c}|^2 \times \sigma^2 + |\triangle\mathbf{c} * \mathbf{r}|^2$, or

$$\approx |\mathbf{c}_0|^2 \times \sigma^2 + |\triangle\mathbf{c}|^2 \times \sigma^2 + |\triangle\mathbf{c} * \mathbf{r}|^2,$$

since the correlation between $\mathbf{c}_0$ and $\triangle\mathbf{c}$ is small. Assuming that the signal-to-noise ratio is large, this becomes

$$\approx |\mathbf{c}_0|^2 \times \sigma^2 + |\triangle\mathbf{c} * \mathbf{r}|^2.$$

Now from (8) and (9) it follows that

$$\triangle\mathbf{c} = \mathrm{IDFT}\left(\frac{1}{R} - \frac{U}{Y}\right)$$

and so

$$|\triangle\mathbf{c} * \mathbf{r}|^2 = \frac{1}{L}\left|\left(\frac{1}{R} - \frac{U}{Y}\right) \times R\right|^2$$

$$= \frac{1}{L}\left|1 - \frac{UR}{UR + W}\right|^2 \approx \frac{1}{L}\left|\frac{W}{UR}\right|^2$$

for $W \ll UR$. Hence,

$$E_1 \triangleq \mathrm{E}(|\triangle\mathbf{c} * \mathbf{r}|^2) = \frac{1}{L}\sum_{i=0}^{L-1} \frac{\mathrm{E}(|W_i|^2)}{|U_i R_i|^2}$$

$$= \sum_{i=0}^{L-1} \frac{\sigma^2}{|U_i R_i|^2}. \tag{11}$$

In the particular case of a channel with small amplitude distortion, that is, with $|R_i| \approx 1$, E_1 is minimized when $|U_i|^2 = L$ for all i, and

$$\min_{\mathbf{u}} E_1 = \frac{1}{L} \sum_{i=0}^{L-1} \frac{\sigma^2}{|R_i|^2} = E_0 .$$

As before, if $\mathbf{u}$ is a maximal-length sequence instead of a zero autocorrelation (ZAC) sequence, E_1 will be multiplied by the factor $2L/L + 1$.

Since the expressions (10) and (11), unlike the expression (6), depend on the R_i, we conclude that statistical knowledge about the channel could be used to reduce the error variance by predistorting the training sequence to compensate for the expected channel characteristics, as indicated in [5]. However, such sequences would not have the ZAC property and are therefore outside the scope of this paper.

Other desirable properties of training sequences

The preceding section demonstrated the advantages of the ZAC property of training sequences for estimating or equalizing an unknown or "reasonably flat" sampled channel. If no restrictions are placed on the values taken by the sequence we obtain an infinity of such sequences for any period L, for it is sufficient to calculate the IDFT of an arbitrary L-sequence of unit phase vectors.

We have already pointed out the desirability of the constant amplitude property. Other desirable properties (depending on the application) are the following:

1. $L = 2^k$, most suitable for fast Fourier transform processing.
2. Sequences taking only real values.
3. Sequences taking a minimum number of values.
4. Sequences taking values in a predetermined signaling set.

Maximal-length sequences satisfy the above properties except for (1) but they do not have the ZAC property. However, they can be modified to acquire the ZAC property by the addition of a suitably chosen constant.

For example let $\mathbf{u}$ be a maximal-length sequence taking the values $+1$ and -1, and let z be a complex constant. The DFT of $\mathbf{u}$ is

$$U_0 = 1, \qquad |U_1|^2 = |U_2|^2 = \cdots = |U_{L-1}|^2 = L + 1.$$

On the other hand, the DFT of the vector $\mathbf{z} = (z, z, \cdots, z)$ is

$$Z_0 = Lz, \qquad Z_1 = Z_2 = \cdots = Z_{L-1} = 0.$$

The DFT of $\mathbf{u} + \mathbf{z}$, which is $U + Z$, will have constant amplitude, and $\mathbf{u} + \mathbf{z}$ will have the ZAC property, if $|1 + Lz|^2 = L + 1$. Hence, any complex constant z satisfying $-1 + 2\mathrm{Re}(z) + L|z|^2 = 0$ will do. The choice of z will therefore also depend on Properties (2), (3), and (4).

For example, if z is chosen to be real,

$$z = \frac{-1 \pm \sqrt{(1 + L)}}{L},$$

then Property (2) will be preserved, but neither constant amplitude (CA) nor Property (4) will hold.

On the other hand, if z is chosen purely imaginary,

$$z^2 = -1/L,$$

then the CA property will be preserved, but not Properties (2) nor (4).

The only known "ideal" CAZAC sequence satisfying all these properties is $+1\ +1\ +1\ -1$.

Unfortunately, it is likely that no other CAZAC sequence taking only the values $+1$ and -1 exists. This has been proven by Turyn [8] for all $L \leq 12\,100$.

The most important class of CAZAC sequences are the so-called *polyphase sequences*, whose elements are roots of unity. It has been shown that such sequences exist for all periods L [9, 10]. If L is odd, an L-phase sequence can be constructed; if L is even, $2L$ phases are needed.

A very important special case consists of sequences of length $n = m^2$, whose elements are mth roots of unity [1, 2], since these satisfy most nearly Properties (3) and (4). Among these, the sequences of length 2^{2k} also satisfy Property (1).

For example, four-phase sequences of length 16, eight-phase sequences of length 64, etc., can easily be constructed. The construction of m^{k+1}-phase sequences of length m^{2k+1} has also been achieved [11]. This allows the construction of four-phase sequences of length 8, eight-phase sequences of length 32, etc., in particular. This construction is described in the Appendix.

The following properties permit the generation of many other sequences, some of which may be preferable to the original constructions.

If u_i is a CAZAC sequence, then so are

- $u_{k \pm i}$, where k is any integer;
- cu_i, where c is any complex constant;
- $u_i W^{ik}$, where k is any integer and W is an nth root of 1;
- $\bar{u}_i$, where $\bar{u}$ denotes complex conjugation; and
- U_i, the DFT of u_i.

Conclusion

CAZAC sequences are useful for channel estimation and fast start-up equalization. They are shown to be optimal under certain assumptions. The most important members of this family are the polyphase sequences of lengths that are a power of 2 because of their suitability for fast Fourier tranform processing.

This paper discusses the use of periodic training sequences for channel estimation and fast start-up equalization, both of which are important in modem applications. A model of a quadrature-amplitude-modulated (QAM) channel is defined. It is then shown that in the presence of noise, error variance of the channel estimate depends on the training sequence used. It is also shown that, under certain assumptions, the error variance is minimized by training sequences possessing the zero-autocorrelation (ZAC) property. Analogous results are obtained for the error variance of equalizer coefficients calculated from a single period of the training sequence. The results obtained with ZAC sequences provide a 3-dB advantage when compared with those obtained with maximal-length sequences.

Other desirable properties of training sequences are further discussed. It is shown that a family of ZAC sequences can be obtained from a maximal-length sequence by the addition of a complex constant. One member of this family, which also has the CA property, is called the CAZAC sequence. Previously published results about polyphase sequences, which are also CAZAC, are summarized, and the construction of a family of m^{k+1}-phase sequences of lengths m^{2k+1} is given in the Appendix. The section ends with a set of properties permitting the generation of related CAZAC sequences. The last of these, which is less well known, may be restated as follows: A sequence is CAZAC if and only if its DFT is CAZAC.

Appendix

The construction of m^{k+1}-phase CAZAC sequences of length $L = m^{2k+1}$ $(k \geq 1)$ is shown below.

The existence of m-phase sequences of length m for m odd and $2m$-phase sequences for m even is known [9].

Let $u_0, u_1, \cdots, u_{m-1}$ be such a sequence. We then form a sequence of length $L = m^{2k+1}$ by enumerating, row by row, the matrix $z_{ij} = u_{i(\bmod m)} W^{ij}$, where

$$i = 0, 1, \cdots, M-1 \qquad M = m^{k+1},$$

$$j = 0, 1, \cdots, N-1 \qquad N = m^k,$$

and W is a primitive Mth root of unity.

This sequence is obviously CA and m^{k+1}-phase; we have to show that it is ZAC.

We observe that the first column of the above matrix is simply the sequence u repeated N times. It is easy to show that its DFT is

$$\underbrace{U_0, 0, 0, \cdots, 0,}_{M-1} \quad \underbrace{U_1, 0, 0, \cdots, 0,}_{M-1} \quad \cdots, \quad \underbrace{U_{m-1}, 0, 0, \cdots, 0,}_{M-1}$$

where U is the DFT of u and the $M-1$ indicates the number of zeros in the row. Also, it is easily shown that the DFT of the nth column is the above sequence rotated n places to the right. Hence the product of the DFTs of the columns is zero, and therefore the columns are orthogonal in pairs. It follows that the autocorrelation coefficients A_l of the sequence obtained by enumerating this matrix row by row are zero for all $l \neq 0 \bmod N$.

It remains to be shown that $A_l = 0$ also when $l = 0 \bmod N$, except when $l = 0 \bmod MN$.

Let $l = cN$ where $c \neq 0 \bmod M$. Then

$$A_l = \sum_{i,j} \bar{z}_{ij} z_{i+c,j},$$

$$= \sum_{i,j} \bar{u}_i W^{-ij} u_{i+c} W^{(i+c)j},$$

$$= \left(\sum_{i=0}^{M-1} \bar{u}_i u_{i+c} \right) \left(\sum_{j=0}^{N-1} W^{cj} \right).$$

We now have two cases to consider:

1. $c \neq 0 \bmod m$; then the first of the above two factors is zero, since u is a ZAC sequence.
2. $c = 0 \bmod m$; then the second factor is zero, since W^c is an Nth root of unity, and $W^c \neq 1$ since $c \neq 0 \bmod M$.

This completes the proof.

References

1. R. C. Heimiller, "Phase Shift Pulse Codes with Good Periodic Correlation Properties," *IRE Trans. Info. Theory* **IT-6,** 254–257 (October 1961).
2. R. L. Frank and S. A. Zadoff, "Phase Shift Pulse Codes with Good Periodic Correlation Properties," *IRE Trans. Info. Theory* **IT-8,** 381–382 (October 1962).
3. W. O. Alltop, "Complex Sequences with Low Periodic Correlations," *IEEE Trans. Info. Theory* **IT-26,** 350–354 (May 1980).
4. R. W. Chang and E. Y. Ho, "On Fast Start-Up Data Communication Systems Using Pseudo-Random Training Sequences," *Bell Syst. Tech. J.* **51,** 2013–2027 (November 1972).
5. K. H. Mueller and D. A. Spaulding, "Cyclic Equalization—A New Rapidly Converging Equalization Technique for Synchronous Data Communication," *Bell Syst. Tech. J.* **54,** 369–406 (February 1975).
6. S. U. H. Qureshi, "Fast Start-Up Equalization with Periodic Training Sequences," *IEEE Trans. Info. Theory* **IT-23,** 553–563 (September 1977).
7. D. N. Godard, "A 9600 Bit/s Modem for Multipoint Communications Systems," *National Telecommunications Conference Record,* New Orleans, Dec. 1981, pp. B3.3.1–B3.3.5.
8. R. Turyn, "Sequences with Small Correlation," *Error Correcting Codes,* H. B. Mann, Ed., John Wiley & Sons, Inc., New York, 1968, pp. 195–228.
9. D. C. Chu, "Polyphase Codes with Good Periodic Correlation Properties," *IEEE Trans. Info. Theory* **IT-18,** 531–532 (July 1972).

10. R. L. Frank, "Comments on Polyphase Codes with Good Periodic Correlation Properties," *IEEE Trans. Info. Theory* **IT-19,** 244 (March 1973).
11. A. Milewski, "Procédé et Dispositif pour Déterminer les Valeurs Initiales des Coefficients d'un Égaliseur Transversal Complexe," French Patent No. 7,540,417, Dec. 30, 1975. A. Milewski, "Method and Apparatus for Determining the Initial Values of the Coefficients of a Complex Transversal Equalizer," U.S. Patent No. 4,089,061, May 9, 1978.

Received Oct. 21, 1982; revised Feb. 18, 1983

Andrzej Milewski *IBM France, Centre d'Etudes et Recherches, 06610 La Gaude, France.* Mr. Milewski is an advisory engineer working on signal processing algorithms with application to modems. He joined IBM in 1961 at Hursley, U.K., and has been at La Gaude since 1962, except for a year's assignment in the San Jose GPD Laboratory in 1981/82. Mr. Milewski received a B.S. in mathematics from the University of the Witwatersrand, Johannesburg, South Africa, in 1954 and an M.A. in mathematics from the University of Cambridge, U.K., in 1956. Before he joined IBM, Mr. Milewski worked for ICL in the U.K. in operations research (1958–61). Since joining IBM, he has worked on programming languages and systems, line switching (network design and performance studies), signal processing with modem applications, multiplexor performance, signal processing with voice applications, and recently, signal processing with applications to magnetic recording. He has published two papers in the *IEEE Transactions on Communications* and he has received two IBM invention achievement awards for his work.