Do the Barker Codes End?
Final Report
MPI 2008, Worcester, MA

1 February 2009

Working Group: *Industrial Representative*: G. Coxson, (Technology Service Corporation); *Faculty*: E. Cumberbatch (Claremont), L.J. Cummings (Nottingham), P. Ferguson (Simon Fraser), I. Mercer (York), C.P. Please (Southampton), and B. Tilley (Olin); *Students*: R. Altalli (Northeastern), L. Cao (NJIT), F. Chen (Purdue), S. Li (UMBC), H. Liang (WPI), Y. Liu (Clemson), J. Miller (Delaware), L. Nguyen (Delaware), M. Salem (Arizona State), P.D. Vu (SUNY Buffalo), J. Watt (Indiana), and Y. Yang (UMBC)

For convenience, we briefly summarize some notation and terminology. A **binary code** is an n-tuple $x = [x_1, \ldots, x_N]$ where $x_i \in S_2 := \{+1, -1\}$ for each i. Then the Cartesian product S_2^N is the set of all 2^N binary codes of length N. For $1 \le k \le N-1$, the kth **autocorrelation** of x is

$$\mathrm{ACF}_x(k) := \sum_{i=1}^{N-k} x_i x_{i+k}.$$

We have $\mathrm{ACF}_x(0) = N$ (the "peak"), and if $k \neq 0$, we refer to $\mathrm{ACF}_x(k)$ as a "sidelobe". Two measures of the collective sidelobe magnitude of a binary code x are the **peak sidelobe level**,

$$\mathrm{PSL}_x = \max_{1 \le k \le N-1} |\mathrm{ACF}_x(k)|,$$

and the **energy**,

$$\mathrm{E}_x = \sum_{k=1}^{N-1} (\mathrm{ACF}_x(k))^2.$$

A **Barker code** is a binary code x with $|\mathrm{ACF}_x(k)| \le 1$ for all $k \neq 0$. If x is a Barker code, then $\mathrm{PSL}_x = 1$, and by a parity argument, $\mathrm{E}_x = \lceil (n-1)/2 \rceil$.

We can also replace the code "alphabet" S_2 with the set of mth roots of unity for some $m > 2$, or the set of all complex numbers of unit modulus. The definition of autocorrelation then changes to $\sum x_i x_{i+k}^*$ where the star denotes complex conjugation. We then call x a **generalized Barker code** if $|\mathrm{ACF}_x(k)| \le 1$ for all $k \neq 0$. Note that in this case, the autocorrelations are no longer restricted to be integers.

At the workshop, the Barker code group split into four non-disjoint subgroups:

- An "algebra group", who explored symmetries of the search space that preserve the autocorrelations' magnitude
- A "computing group", who explored methods for quickly finding binary codes with very good autocorrelation properties
- A "statistics group", who explored ways to quantify what has been empirically observed about autocorrelation in the search space S_2^N
- A "continuous group", who explored a non-discrete analogue of the problem of finding sequences with good autocorrelations.

What follows is a summary of what each of the four subgroups were able to accomplish at the workshop.

1 Algebraic structure

Both in the case of binary sequences and sequences consisting of mth roots of unity, there are certain operations that preserve the absolute value of the autocorrelations and hence preserve the PSL and the energy. Note that our "search space" grows exponentially (there are 2^N binary sequences of length N, and m^N sequences of length N consisting of mth roots of unity). There are some methods known for finding the best sequence of length N that improve somewhat upon naive brute force search, by taking advantage of symmetries of the search space as well as using techniques such as "branch and bound" search. For example, some known algorithms for finding the minimum energy or minimum PSL in the binary case have running times of roughly 1.85^N and 1.4^N respectively, as opposed to 2^N. See [3, 4] and [1, 2] respectively. Further understanding of the symmetries of the search space could result in further reductions in search time, both in the binary and nonbinary case. For example, the problem presenter mentioned at the workshop that it was not known to him whether the structure of the group of PSL-preserving symmetries is known in the case of mth roots of unity for even length N.

2 Computational exploration

It can be valuable not only to develop exhaustive algorithms that find the minimum PSL or energy among all binary sequences of length N, but also to find practical methods to quickly find sequences with close to optimal PSL or energy. Often in real-world applications, it suffices to have sequences with good PSL or energy, as opposed to best possible PSL or energy among all sequences. Thus, one is sometimes willing to sacrifice exhaustiveness if one can find good sequences quickly.

There are theoretical results giving necessary conditions on the ratio between the number of positive entries and negative entries in a binary Barker code (if they exist). We can also observe the ratio of positive to negative entries in tables of known sequences with optimum or near-optimum PSL or energy, and then decide to search through only those sequences having a certain ratio of positive to negative entries.

At the workshop, we experimented with such searches. Moreover, rather than performing an exhaustive search of all binary sequences with a particular ratio of positive to negative entries, we attempted to conduct more "intelligent" searches using heuristic methods to quickly "zero in" on good sequences, by setting a particular ratio of positive to negative entries at the outset and then permuting entries.

3 Statistical analysis

For values of N up to a certain point, it is possible to exhaustively compute the energy for all binary sequences of length N. These lengths can be plotted in a histogram. When we do so, a pattern is immediately clear. We observed during the workshop that these histograms conform astonishly well to a well-known probability distribution called the Gumbel distribution or extreme value distribution.

In fact, the data for energy of binary sequences fits this probability distribution so well that by the standards of empirical science, there is no practical doubt that it is the "correct" distribution. However, providing a rigorous analytic proof of this statement may be difficult.

The left tail of the Gumbel distribution approaches 0 like $e^{-e^{-x}}$. One could prove analytically that there are no Barker codes of length N if one could show that the probability of a binary code having energy $\lceil (N-1)/2 \rceil$ is bounded above by some quantity less than 2^{-N} (since there are 2^N binary sequences of length N). The rapid decrease of the left tail of the Gumbel distribution, together with the excellence of the fit of the distribution to the data, is strong evidence for the nonexistence of Barker codes for large N, although an analytic proof continues to elude us.

4 Continuum model

The connection between communications systems and binary codes dates back to the middle of the 20th century. Since most communications systems rely on wave propagation through a continuous medium, the means by which information is transmitted is based on the solutions of linear partial differential equations. These equations, which rely on a balance between spatial and temporal gradients, are solved using the classical techniques whose basis is found in Fourier analysis. Further, since the original communications problems were based on electrical signals, the fundamental spatial component of the equation could be represented as a known quantity (the carrier wave), and the resulting mathematical problems for information transmission depend on the modulation of the signal, either through amplitude modulation or through phase modulation. These classical techniques can be found in any text in continuous and discrete signals processing, such as Oppenheim et al. [5].

The connection between the continuous Fourier representation described above and the discrete binary sequences discussed in the workshop has been outlined in [6]. We outline the discussion here based on the continuum approach used at the workshop. Consider the real-valued signal

$$f(t^*) = A(t^*/T_1)e^{i[\phi(t^*/T_2)]} + c.c$$

where t^* is the dimensional time of interest, A is the real amplitude modulation, a function that is $2\pi T_1$-periodic, and ϕ is the phase modulation that is $2\pi T_2$-periodic. If this signal is sent out from a transmitter, then an example of a received signal is given by g,

$$g(t^*) = f(t^* + \tau^*) = A((t^* + \tau_s^*)/T_1)e^{i\,\phi((t^*+\tau_s^*)/T_2)} + c.c$$

where τ_s^* is the time shift of interest. The autocorrelation function of f is given by

$$I(\tau^* - \tau_s^*) = \frac{1}{2\pi T_1}\int_{-\pi T_1}^{\pi T_1} f(t^*)\bar{f}(t^* + \tau^* - \tau_s^*)\,dt^* \tag{1}$$

where I is the continuous form of the autocorrelation function that is desired. For simplicity, let us scale time $t^* = T_1 t$, where t is now a dimensionless time to arrive at the autocorrelation function

$$I(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} A(t)A(t+\tau)e^{i\,[\phi(\ell\, t)-\phi(\ell(t+\tau))]}\,dt \tag{2}$$

where $\ell = T_1/T_2 \in \mathbf{Z}$ is the ratio of the amplitude modulation period to the phase modulation period, and we have centered τ to the desired time shift

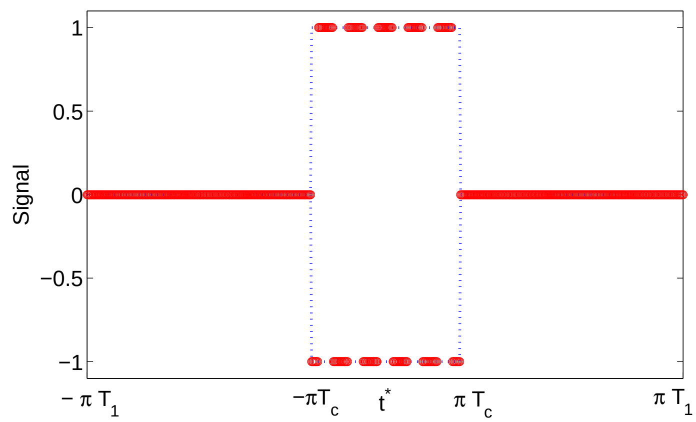

Figure 1: Cartoon of sample signal along one period. The red dots correspond to the actual signal, while the dashed line corresponds to the amplitude modulation in t^*, which has an extent T_c/T_1 of the fundamental period. The phase modulation ϕ is conveyed by the change in sign of the signal with its own periodicity of $2\pi T_2$ (not labelled).

(i.e. $\tau = 0$ corresponds to the proper return signal time). In the classical connection of discrete codes to the Fourier context, $A(t) = p_\mu(t)$, where $\mu = T_c/T_1$ is the fraction of time when $A \neq 0$, and $\ell = 1$ (see [6]). A pictorial representation of this setup is shown in Figure 1. In the problem presented by TSC at the workshop, ℓ is the code length.

The goal is to determine the phase ϕ and the amplitude A such that the difference between the convolution I and a desired autocorrelation function is minimized. The ideal autocorrelation function would be unity at the signal return time and zero at other times. This would correspond to an infinite peak-to-peak amplitude ratio, since the off-peak amplitudes are zero. We can define this function $\sigma(\tau)$ in terms of a Cauchy sequence

$$\sigma(\tau) = \lim_{\mu\to 0} \lim_{n\to\infty} I_n(\tau) = \lim_{\mu\to 0} \lim_{n\to\infty} \left(\frac{\sin(\tau/\mu)}{\tau/\mu} \right)^{2n} .$$

A sampling of these I_n is shown in Figure 2. Note that $I_{1/2}$ corresponds to the inverse Fourier transform of the spectrum $p_{1/\mu}$ for μ fixed. Although this function can be approximated by many Cauchy sequences, for the purposes of this problem this choice is instructive, since the peak-to-peak amplitude of I_n decreases with increasing n.

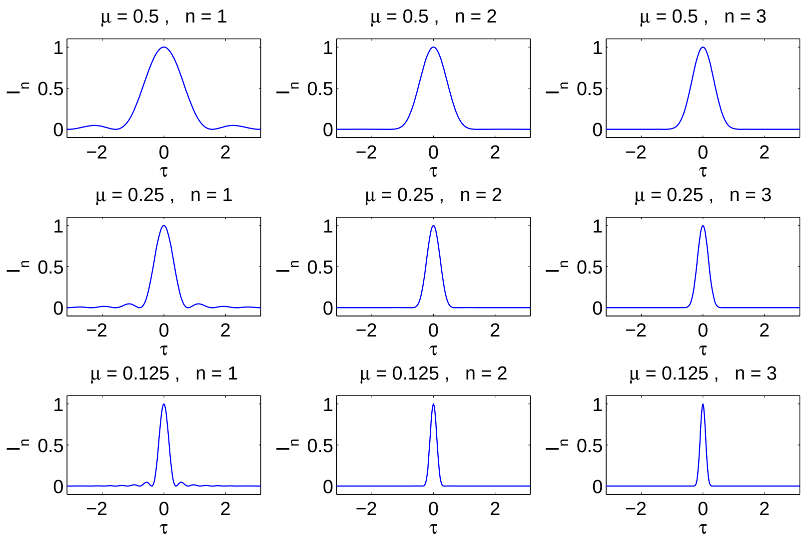

Figure 2: Sample autocorrelation functions I_n for a range of exponents n and periodic box sizes L. $N = 256$ collocation points were used in the discretization. Note that the peak-to-peak ratio increases with the exponent n

To find A, ϕ, we take advantage of properties of Fourier transforms. Note that I_n is an even function of τ, and the integral (2) is a convolution integral for $\tau \to -\tau$. Hence if $\mathcal{F}$ denotes the Fourier transform operator in τ, then

$$\mathcal{F}\{I_n\} = \mathcal{F}\{f * f\} = [\mathcal{F}\{f\}]^2 \ .$$

Hence, the solution to f is given by

$$f = \mathcal{F}^{-1}\left\{\sqrt{\mathcal{F}\{I_n\}}\right\} \ , \tag{3}$$

provided that $I_n(-\tau) = I_n(\tau)$. Note that we have chosen the positive signed square root. Since the negative branch is also a solution, the transformation $\phi \to \phi \pm \pi$ results in solutions corresponding to this branch.

To find f based on these I_n, we then apply the Fast Fourier Transform using Matlab to perform the forward and inverse Fourier transforms. Note that this computational technique requires a maximum number of collocation points at which to evaluate the autocorrelation function I_n. The number of these points N provides an upper limit for the length of the code of interest (i.e. $\ell \leq N$). We anticipate that the length of the code is to be determined by the solution. All of the results presented below have been verified to satisfy the autocorrelation relation to within roundoff error.

Figure 3 shows the amplitude modulation A for the same autocorrelation functions given in Figure 2. Notice that the amplitude modulation is typically localized near the center $\tau = 0$.

Figure 4 shows the phase modulation ϕ given by the solution for the same autocorrelation functions shown in Figure 2. Notice that the phase shifts are binary: either $\phi = 0$ or $\phi = \pm\pi$. Since the Barker code corresponds to $\cos\phi$, the sign of the $\phi = \pm\pi$ cases is irrelevant. Figure 5 shows the Barker code for each of these autocorrelation functions. Although the mapping is not known between these codes and what the equivalent binary Barker code would be, it appears that each of these cases could be construed as a simple binary Barker code $[1, -1]$ or $[-1, 1]$.

As a final note, if we increase the size of n and decrease the size of μ, we can find codes that do not correspond clearly with a binary Barker code. Figure 6 shows a situation where the code appears to deviate significantly from one case to a second. A close-in look at this last case (see Figure 7a) shows that this sequence does not match cleanly with any of the known Barker codes. Figure 7b shows the amplitude modulation A for this particular solution. Notice that the resolution of the amplitude modulation is poor, but the choice of the phase modulation corrects this to produce an accurate autocorrelation function.

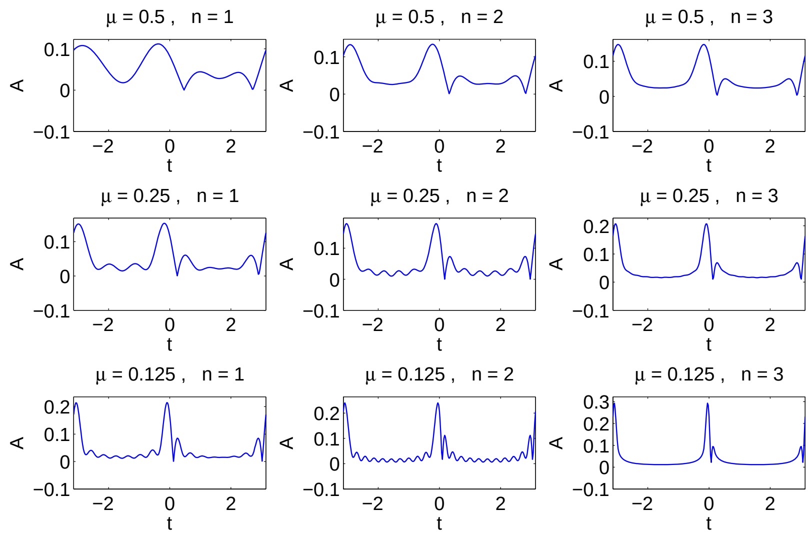

Figure 3: Amplitude modulation of f for a given autocorrelation function shown in Figure 2.

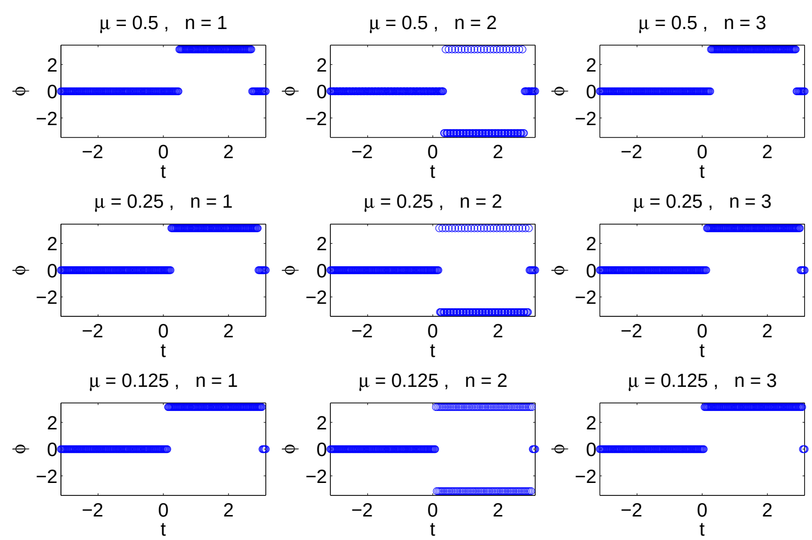

Figure 4: Phase modulation ϕ of f for a given autocorrelation function shown in Figure 2

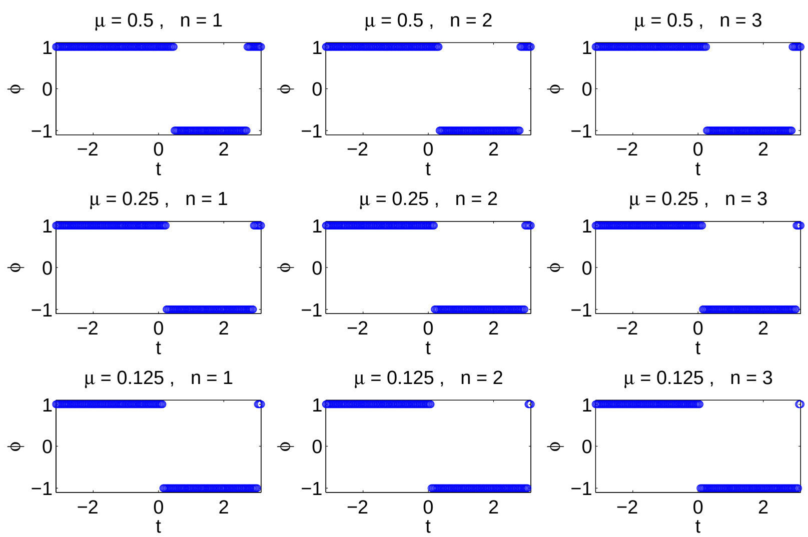

Figure 5: Barker code representation $\cos\phi$ based on the phase modulation results shown in Figure 4. Notice that the mapping of these codes to the binary Barker codes is an open question.

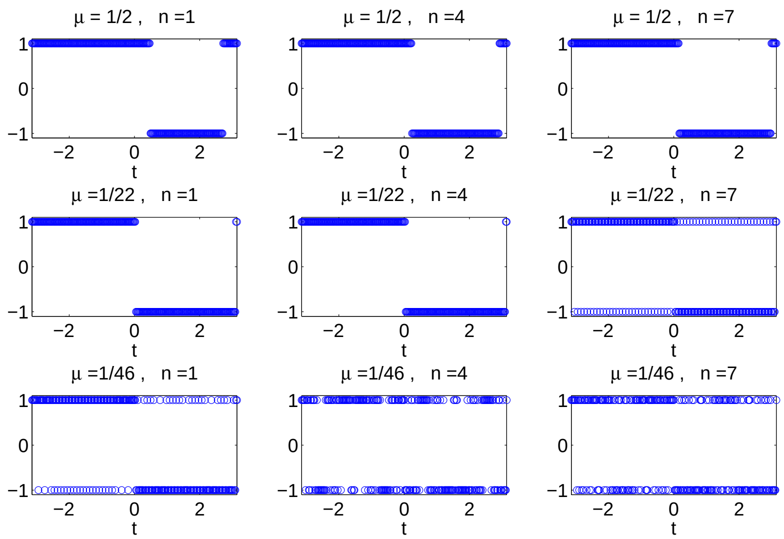

Figure 6: Note that the increasing values n and decreasing values of μ can result in nontrivial Barker codes, as noted in the lower right figure.

In conclusion, we have found through a direct Fourier representation a way to generate codes which depend on both amplitude and on phase. The classical formulation of the code assumes *ab initio* that the amplitude modulation is piecewise constant. Were one to impose this amplitude restriction, the set of phases ϕ that would best mimic the autocorrelation function for amplitude modulation A is the solution of a constrained optimization problem. It is possible that a formulation based on the Fourier representation could be implemented to provide insight into the existence of different codes, not necessarily binary Barker codes, that may be useful for TSC to consider for its applications.

References

[1] M.N. Cohen, M.R. Fox and J.M. Baden, "Minimum peak sidelobe pulse compression codes", in *IEEE International Radar Conference*, pages 633–638. IEEE, 1990.

[2] G.E. Coxson, A. Hirschel and M.N. Cohen, "New results on minimum-PSL binary codes", in *IEEE Radar Conference*, pages 153–156, IEEE,

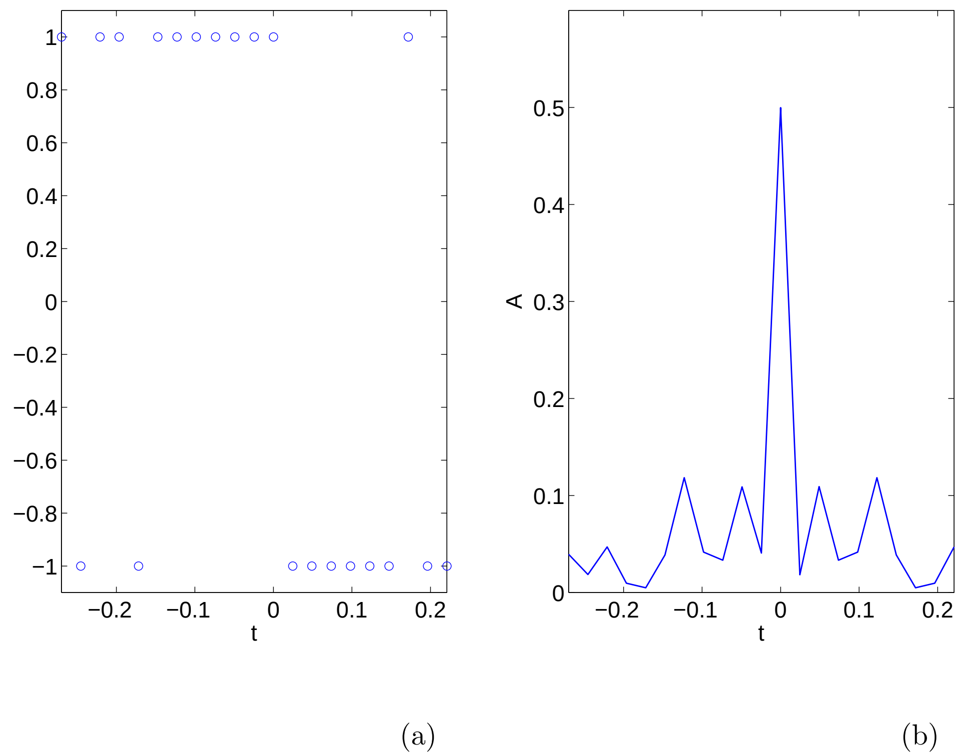

Figure 7: (a) A close-in look at the code for $\mu = 1/46, n = 7$ with $N = 256$. Notice that the sequence of the code does not resemble the known binary Barker codes. (b) The corresponding amplitude modulation A.

2001.

[3] S. Mertens, “Exhaustive search for low-autocorrelation binary sequences”, *J. Phys. A: Math. Gen.*, **29**: L473–L481 (1996).

[4] S. Mertens and H. Bauke, “Ground states of the Bernasconi model with open boundary conditions”, available at <http://odysseus.nat.uni-magdeburg.de/∼mertens/bernasconi/open.dat> (2004).

[5] A.V. Oppenheim and A.S. Willsky and S.H. Nawab, *Signals and Systems*, 2nd edition, Prentice-Hall (1997).

[6] D.V. Sarwate and M.B. Pursley, “Crosscorrelation properties of pseudo-random and related sequences”, *Proc. IEEE*, **68** (5) 593-619: (1980).