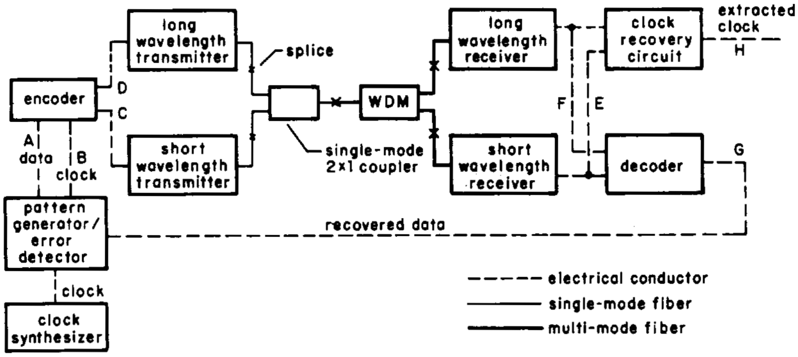
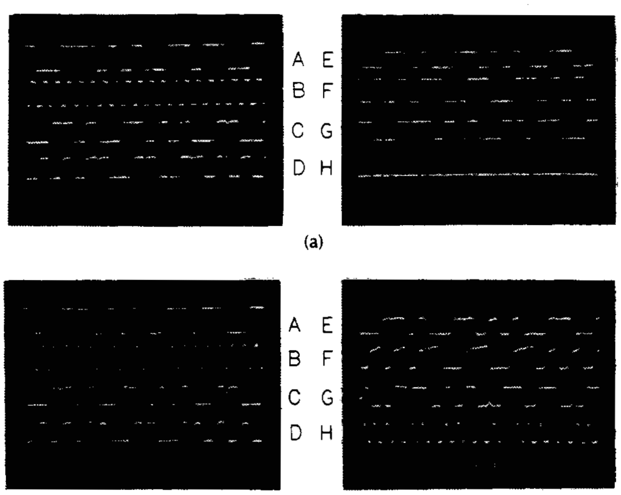


Fig. 5. Block diagram of experimental setup for code demonstration.

end by introducing electronic or fiber delay on the faster channel (longer wavelength). In the system under consideration (taking the propagation delay difference to be 0.1 ns per kilometer of fiber per nanometer of wavelength difference) a 25-ns/km delay difference occurs. The delay difference could be substantially decreased if the two wavelengths were closer to one another, yet distinguishable by a wavelength demultiplexer. Typically, a commercial WDM can separate wavelengths 5 percent apart. Thus if $S1$ and $S2$ were transmitted at 1300 and 1365 nm, respectively, the propagation delay difference would be reduced to 6.5 ns/km.

Fig. 6(a) shows: a 1-Mbit/s random data sequence A; a 1-MHz clock B; encoded waveforms C and D before optical conversion; encoded sequences E and F after transmission; decoded data G; and extracted clock H. Measurement points A through H are indicated in Fig. 5. The input data rate then is changed from 1 to 10 Mbits/s, and the waveforms at the same measurement points A through H are shown in Fig. 6(b). The clock and data are again recovered, with no modifications necessary to the encoder, decoder, or clock extraction circuits. No transmission errors were recorded in 10^{10} clock cycles (BER of 10^{-9}) at either transmission rate.



(b)

Fig. 6. Waveforms at characteristic points of the system configuration for (a) 1 Mbit/s and (b) 10 Mbits/s. A—input data waveform to be encoded; B—original clock waveform; C—short-wavelength channel encoded sequence; D—long-wavelength channel encoded sequence; E—detected short-wavelength channel sequence; F—detected long-wavelength channel sequence; G—decoded data waveform; H—extracted clock waveform.

In summary, a new digital line code, named alternate mark/space inversion (AMSI) code, is demonstrated. It has zero dc content, conserves baseband bandwidth, is self-clocking, and is transparent to the data rate. Since the original data bandwidth is maintained, the cost of implementing the two required channels might compare favorably to the cost of implementing one higher bandwidth channel (as is usually required for a conventional code). A unique feature of the AMSI code is that both data and clock can be extracted digitally. This property implies that the data transmission rate can be changed without affecting the encoding, decoding, and clock extraction process, as long as the hardware speed is not exceeded. Such a rate-transparent system is useful in applications where the anticipated data rate is unknown, or when different rate services share the same system.

ACKNOWLEDGMENT

The authors would like to thank C. Brandmaier for her assistance in designing the encoder and decoder circuits and M. Chan for her contribution in the clock extractor circuit design.

REFERENCES

[1] K. W. Cattermole, "Principles of digital line coding," *Int. J. Electron.*, vol. 55, no. 1, pp. 3-33, 1983.
[2] Y. Takasaki, M. Tanaka, N. Maeda, K. Yamashita, and K. Nagano, "Optical pulse formats for fiber optic digital communications," *IEEE Trans. Commun.*, vol. COM-24, no. 4, pp. 404-413, Apr. 1976.
[3] J. C. Mallinson and J. W. Miller, "Optimal codes for digital magnetic recording," *Radio Electron. Eng.*, vol. 47, no. 4, pp. 172-176, Apr. 1977.
[4] W. E. Little, and H. K. Clark, "PCM and data transmission codes," *Telephone Eng. Manag.*, vol. 82, no. 15, pp. 84-8, 90, Aug. 1, 1978.
[5] P. R. Prucnal, P. A. Perrier, and C. Brandmaier, "A low dc, self-clocking, two-color transmission code for optical communications," *Proc. IEEE*, vol. 75, no. 6, pp. 856-858, June 1987.
[6] J. A. Guinea, C. Stacey, and R. Batruni, "Digital transmission in the subscriber loop," *IEEE Circuits Dev. Mag.*, vol. 2, no. 5, pp. 14-28, Sept. 1986.

On Predicting a Band-Limited Signal Based on Past Sample Values

P. P. VAIDYANATHAN

The prediction of samples of a band-limited real signal $x_a(t)$ from a finite number of past samples is considered. It is shown how a signal-independent linear predictor of finite order can be constructed based on Chebyshev polynomials, such that the prediction error tends to zero for sampling rate exceeding the Nyquist rate.

Manuscript received December 29, 1986. This work was supported by the Naval Oceans Systems Center under Contract N66001-85-D-0203, through a subcontract by the San Diego State University Foundation, and by Caltech's programs in Advanced Technology, sponsored by Aerojet General, General Motors, GTE, and TRW.
The author is with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, USA.
IEEE Log Number 8714507.

0018-9219/87/0800-1125$01.00 © 1987 IEEE

I. Introduction

An interesting problem in linear-prediction theory is the following: Let $x_a(t)$ be a real continuous-time signal, band-limited to the region $|\Omega| \le \Omega_M$. What is the smallest sampling frequency Ω_S which will enable us to predict the present sample values $x_a(nT_S)$ based on a finite number of past samples, with arbitrarily small (prespecified) error, and with predictor coefficients *independent of* the signal $x_a(t)$? (Here T_S represents the sampling period, i.e., $T_S = 2\pi/\Omega_S$.) In a 1962 text [1, pp. 70–73], Wainstein and Zubakov showed that such prediction is possible as long as the sampling frequency Ω_S satisfies $\Omega_S > 3\Theta$ where Θ is the Nyquist frequency given by $\Theta = 2\Omega_M$. Explicit formulas for predictor coefficients $a_{N-1,k}$ of the $(N-1)$th order predictor (where N depends on the desired prediction accuracy) are also given in [1]. As the predictor order $(N-1) \to \infty$, an appropriate norm of the prediction error approaches zero.

In 1972, Brown, Jr., [2] extended these results and showed that it is *sufficient* to sample the signal $x_a(t)$ at *two* times the Nyquist frequency Θ, and also showed how to obtain the coefficients of the $(N-1)$th order predictor. These results were further improved by Splettstosser [3] in 1982, who showed that this kind of a prediction is possible even with the sampling frequency equal to 1.5 times the Nyquist frequency Θ. Once again, the coefficients $a_{N-1,k}$ of the $(N-1)$th order predictor (which were independent of $x_a(t)$) could be derived.

Brown [2] and Splettstosser [3] have also observed that it is theoretically possible to predict the samples of $x_a(t)$ in the above manner, as long as the sampling frequency is larger that the Nyquist rate by any arbitrarily small amount $\epsilon > 0$, i.e.,

$$\Omega_S = \Theta + \epsilon = 2\Omega_M + \epsilon. \tag{1}$$

(However, explicit expressions for the predictor coefficients $a_{N-1,k}$ in this case could not be obtained.) This observation has also been recently made by Papoulis [8] who has given a different proof. Further references, and proofs can be found in [9], [10] and the references contained therein. The purpose of this letter is to give a *constructive* proof that (1) is sufficient. The proof is based on the use of the Chebyshev polynomial, and automatically shows how the predictor coefficients can be computed (once again, independent of $x_a(t)$).

II. Construction of the Predictor Coefficients

In order to keep the notations simple, we shall present our arguments based on the sequence $x(n)$ defined by $x(n) = x_a(nT_S)$ and its transform [4] given by

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x(n)\, e^{-j\omega n}.$$

Let $x(n)$ be band-limited to $|\omega| \le \omega_M < \pi$ (see Fig. 1). We can re-cast the prediction problem by asking how large ω_M can be, so that $x(n)$

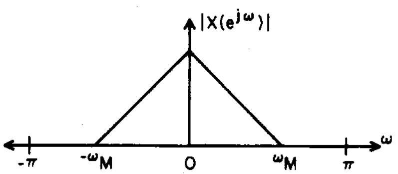


Fig. 1. The band-limited spectrum of $x(n)$.

can *still* be predicted from a finite number of its past values $x(n-k)$, $k > 0$ with arbitrary accuracy. The answer, based on the results in [2], [3], [8], [9] is, of course, that ω_M can be of any value less than π. We shall present a new method for constructing the predictor coefficients.

The basic principle behind the construction of the predictor coefficients in [1]–[3] can be summarized as follows: let $H(z)$ be the transfer function of a finite impulse response (FIR) filter of order $L-1$ given by

$$H(z) = \sum_{k=0}^{L-1} h(k)\, z^{-k}.$$

Here $h(n)$ is the impulse response of $H(z)$. Suppose $H(z)$ is a high-pass filter with stop-band edge ω_M, so that $|H(e^{j\omega})|$ is "small" in $0 \le |\omega| \le \omega_M$ (see Fig. 2(a)). If $x(n)$ is passed through this filter, the output sequence $e_h(n)$ therefore has "small" energy. We have

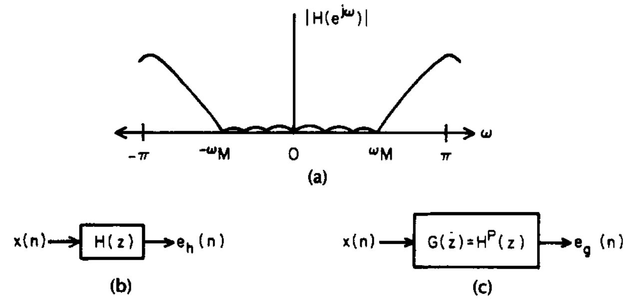


Fig. 2. High-pass filtering of $x(n)$.

$$0 \simeq e_h(n) = \sum_{k=0}^{L-1} h(k)\, x(n-k).$$

If we assume that $h(n)$ is normalized such that $h(0) = 1$, then this implies

$$x(n) \simeq \hat{x}(n) = -\sum_{k=1}^{L-1} h(k)\, x(n-k).$$

In other words, as long as $H(z)$ is a good high-pass filter with $h(0) = 1$ and with stop-band edge $\ge \omega_M$, the coefficients $-h(1), -h(2), \cdots, -h(L-1)$ can be taken as the coefficients of the linear predictor, i.e., $a_{L-1,n} = -h(n)$, $1 \le n \le L-1$.

If the above high-pass filter $H(z)$ is not a "good" filter in the sense that its stop-band attenuation is not large enough, then a new high-pass filter $G(z) = H^P(z)$ can be used for the construction of the predictor, as long as $|H(e^{j\omega})| \le H_{\max} < 1$ for $|\omega| \le \omega_M$. Here, P is a positive integer, and the FIR filter

$$G(z) = g(0) + g(1)\, z^{-1} + \cdots + g(N-1)\, z^{-(N-1)}$$

has $N-1 = P(L-1)$, and $g(0) = 1$ since $h(0) = 1$. Its stop-band attenuation can be made arbitrarily large by increasing P. As a result, with $a_{N-1,n} = -g(n)$, $1 \le n \le N-1$, we have the prediction error $e_g(n) \to 0$ as $P \to \infty$ (see Fig. 2(c)).

For example, in [1] we have $H(z) = 1 - z^{-1}$, and $G(z) = H^P(z)$. Clearly, $|H(e^{j\omega})| = |2\sin(\omega/2)|$ is less that unity when $\omega_M < 2\pi/6$. In other words, the sampling frequency 2π must be greater than three times the Nyquist rate $2\omega_M$, if this cohice of $H(z)$ has to work. In order to permit values of ω_M larger than $2\pi/6$, we can choose $H(z)$ differently; the only constraint on $H(z)$ is that it should be FIR with $h(0) = 1$ and satisfy $|H(e^{j\omega})| < 1$ for $|\omega| \le \omega_M$.

Our construction of $H(z)$ here is based on the Dolph-Chebyshev window function [5] $D_K(\omega)$ which is defined as

$$D_K(\omega) = C_K\left(\cos\frac{\omega}{2}\Big/\cos\frac{\omega_c}{2}\right)$$

where $C_K(x)$ is the Chebyshev polynomial [6, pp. 376–380] of degree K, in the variable x. We assume K is even, so that

$$C_K(x) = \sum_{n=0}^{K/2} c_{2n} x^{2n}.$$

A plot of $D_K(\omega)$ is shown in Fig. 3. The significant feature here is that, it has a low-pass spectrum with equiripple stop-band attenuation. The stop-band, which covers the region $\omega_c \le \omega \le \pi$ has peak ripple

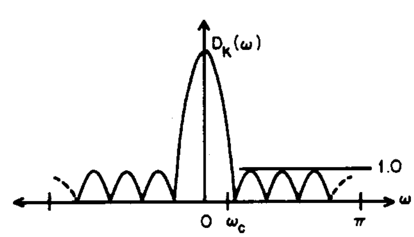


Fig. 3. Plot of the Dolph-Chebyshev response.

equal to unity. If we now replace $\cos(\omega/2)$ in $D_K(\omega)$ with $(z^{1/2} + z^{-1/2})/2$ and multiply the result by $z^{-K/2}$ we obtain a causal low-pass FIR transfer function

$$F(z) = \sum_{n=0}^{L-1} f(n) z^{-n}$$

of order $L - 1 = K$, with

$$|F(e^{j\omega})| = |D_K(\omega)|.$$

An FIR high-pass filter of order $L - 1$ with stop-band edge $\pi - \omega_c$ can be obtained from here by defining it to be $F(-z)$. It can be verified that the impulse response coefficient $f(0)$ of $F(z)$ is given by

$$f(0) = c_K \Big/ \left(2^K \cos^K \left(\frac{\omega_c}{2}\right)\right)$$

where c_K is the coefficient of $C_K(x)$ corresponding to x^K. It is well known [6] that $c_K = 2^{K-1}$, hence we have

$$f(0) = 1/[2 \cos^K (\omega_c/2)].$$

Accordingly, if we define a high-pass filter $H(z)$ to be

$$H(z) = 2 \cos^K \left(\frac{\omega_c}{2}\right) F(-z) = \sum_{n=0}^{L-1} h(n) z^{-n} \tag{2}$$

then $H(z)$ has the following features:

1) $h(0) = 1$.
2) $|H(e^{j\omega})| \le 2 \cos^K (\omega_c/2)$ for all ω such that $|\omega| \le \pi - \omega_c$.

Now suppose we are given the band-limited signal $x(n)$, with $\omega_M = \pi - \epsilon$, where $\epsilon > 0$. We then choose $\omega_c = \epsilon$ so that the stop-band of $H(z)$ coincides with the signal band. A large enough even-integer K can now be found such that $2 \cos^K (\epsilon/2) < 1$. Such a finite K exists because $\epsilon > 0$. Once ω_c and K are thus found, $H(z)$ is known. We now define $G(z) = H^P(z)$. As P increases, the stop-band energy of $H^P(z)$ gets smaller and smaller, and the prediction accuracy increases, i.e., $e_g(n) \to 0$ as $P \to \infty$. (Formally, appropriate norms for the error can be formulated depending on whether the signal is deterministic or stochastic [1]-[3]. These norms tend to zero as $P \to \infty$.) The impulse response coefficients $g(n)$ of $G(z)$ directly give us the predictor coefficients $a_{N-1,n} = -g(n)$, $1 \le n \le N - 1$ where $N - 1 = P(L - 1) = PK$.

III. Practical Considerations

Some of the coefficients $a_{N-1,n}$ computed as above take on very large values, as N increases. This feature is also true of the predictor coefficients in [1], which are based on binomial coefficients. Brown [2] has modified the results in [1] by incorporating a factor α such that $0 < \alpha < 1$. According to Brown's method we have $H(z) = (1 - \alpha z^{-1})$, and $G(z) = H^P(z)$, so that the predictor coefficients are

$$a_{N-1,n} = -(-\alpha)^n \binom{N-1}{n}, \quad 1 \le n \le N - 1$$

where $N - 1 = P$. The presence of α here helps to reduce the growth of the numbers $a_{N-1,n}$ to some extent. As ω_M gets closer to $\pi/2$, α gets smaller. But at the same time, $|H(e^{j\omega_M})|$ gets closer to unity. As a result, P (and hence N) has to be made very large in order to reduce the prediction error, whenever α is small. The net effect on the size of the coefficients $a_{N-1,k}$ is that, they continue to be large. Accordingly, in a practical implementation (say, with fixed-point arithmetic) of predictors based on most of these techniques, one has to take into account the possibility of internal computational overflow and also roundoff noise.

References

[1] L. A. Wainstein and V. D. Zubakov, *Extraction of Signals from Noise.* London, England: Prentice-Hall Int., 1962.
[2] J. L. Brown, Jr., "Uniform linear prediction of band-limited processes from past samples," *IEEE Trans. Informat. Theory*, vol. IT-18, pp. 662-664, Sept. 1972.
[3] W. Splettstosser, "On the prediction of band-limited signals from past samples," *Informat. Sci.*, vol. 28, pp. 115-130, 1982.
[4] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1975.
[5] H. D. Helms, "Digital filters with equiripple or minimax responses," *IEEE Trans. Audio Electroacoust.*, vol. AU-19, pp. 87-94, Mar. 1971.
[6] M. E. Van Valkenburg, *Modern Network Synthesis.* New York, NY: Wiley, 1964.
[7] D. H. Mugler and W. Splettstosser, "Difference methods and roundoff-error bounds for the prediction of band-limited functions from past samples," *Frequenz*, vol. 39, pp. 182-187, July-Aug. 1985.
[8] A. Papoulis, "A note on the predictability of band-limited processes," *Proc. IEEE*, vol. 73, pp. 1332-1333, Aug. 1985.
[9] F. Marvasti, "Comments on '*A note on the predictability of band-limited processes*'," *Proc. IEEE*, vol. 74, p. 1596, Nov. 1986.
[10] J. L. Brown, Jr. "On the prediction of a band-limited signal from past samples," *Proc. IEEE*, vol. 74, pp. 1596-1598, Nov. 1986.

A Separable Hartley-Like Transform in Two or More Dimensions

MICHAEL G. PERKINS

This letter introduces a discrete, separable, real-to-real transform, called the cas-cas transform. Theorems for the two-dimensional (2-D) case are presented, and the cas-cas transform is compared to the Hartley transform as an alternative way to convolve 2-D arrays and compute 2-D power spectra.

I. Introduction

The n-dimensional Discrete Hartley Transform (DHT) is defined by

$$H(u_1, u_2, \cdots, u_n) = \sum_{m_1=0}^{M_1-1} \sum_{m_2=0}^{M_2-1} \cdots \sum_{m_n=0}^{M_n-1} f(m_1, m_2, \cdots, m_n) \cdot \operatorname{cas}\left(2\pi \sum_{i=1}^{n} u_i m_i / M_i\right)$$

where $\operatorname{cas} \alpha = \cos \alpha + \sin \alpha$ [1]. The DHT is a real-to-real transform, and can therefore be computed without the complex arithmetic necessary to compute the DFT. In addition, the DHT uses the same kernel for both the forward transformation and the inverse transformation, so that

$$f(m_1, m_2, \cdots, m_n) = \frac{1}{\prod_{i=1}^{n} M_i} \cdot \sum_{u_1=0}^{M_1-1} \sum_{u_2=0}^{M_2-1} \cdots \sum_{u_n=0}^{M_n-1} H(u_1, u_2, \cdots, u_n) \cdot \operatorname{cas}\left(2\pi \sum_{i=1}^{n} u_i m_i / M_i\right).$$

Fast algorithms for computing one-, two-, and three-dimensional DHTs have been published [2]-[4].

The n-dimensional Hartley transform is not a separable transform, however, because for $n > 1$,

$$\operatorname{cas}\left(\sum_{i=1}^{n} A_i\right) \ne \prod_{i=1}^{n} \operatorname{cas}(A_i).$$

For some applications, a separable transform is desirable. An example is found in transform image coding where, if a separable trans-

Manuscript received December 1, 1986. This work was supported by NASA under Contract NAGW-419.

The author is with the Center for Aeronautics and Space Information Sciences, Stanford University, Stanford, CA 94305-4055, USA.

IEEE Log Number 8714174.

0018-9219/87/0800-1127$01.00 © 1987 IEEE