This is to certify that the

dissertation entitled

Adaptive Filtering in Remote Heart Rate Measurements

presented by

William J. Byrne III

has been accepted towards fulfillment
of the requirements for

M. S. degree in Electrical Engineering

Marvin Siegel
Major professor
Marvin Siegel

Date 6-19-85

ADAPTIVE FILTERING IN REMOTE HEART RATE MEASUREMENTS

By

William Joseph Byrne III

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Electrical Engineering and Systems Science

1985

ABSTRACT

ADAPTIVE FILTERING IN REMOTE HEART RATE MEASUREMENTS

By

William Joseph Byrne III

Recent efforts have been directed towards developing bio-medical instrumentation which measures human heart rate using microwave energy. A low-level microwave signal is transmitted and received by a portable, self-contained homodyne transceiver system. The instrument is either placed directly on the subject's chest or pointed at the chest from a distance of several feet. The chest is illuminated with microwave energy and doppler shifts in the reflected signal are used to measure chest motion. This motion contains components due to breathing and the heart beat as well as clutter components due to upper body movement and channel noise obscure the heart signal.

This presentation will explain the applicability of adaptive filtering to the problem of microwave heart rate measurements. The derivations of several adaptive filters will be presented. A discussion of results will be presented which will compare the behavior of several adaptive algorithms and measure their performance.

ACKNOWLEDGMENTS

I would like to express my indebtedness to my major advisor, Dr. Marvin Siegel, for permitting me the opportunity to study this material. I am also most grateful to Dr. H. Roland Zapp, who spent much time reviewing my work. I also appreciate the assistance of my fellow students, Gregory Hoshal, who offered many helpful suggestions based on his experience in this area, and Patrick Flynn, whose laboratory computing support made my work much easier.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Recent efforts have been directed towards developing bio-medical instrumentation which measures human heart rate using microwave energy. A low-level microwave signal is transmitted and received by a portable, self-contained homodyne transceiver system. The instrument is either placed directly on the subject's chest or pointed at the chest from a distance of several feet. The chest is illuminated with microwave energy and doppler shifts in the reflected signal are used to measure chest motion. This motion contains components due to breathing and the heart beat as well as clutter components due to upper body movement and channel noise which act to obscure the heart signal. If the occurrence of the heart beats can be accurately determined it is possible to estimate the heart rate. This task is complicated by the clutter in the signal. Previous detection schemes have used peak detection to detect heart beats or autocorrelation to estimate the heart rate, with limited success.

Peak detection requires that the signal be fairly uncluttered. There should be no constructive interference from noise breathing or background motion which might cause false peaks in the signal. Peak detection works well when the subject is rested and breathing regularly and the transceiver is not too far from the subject. In other circumstances, peak detection performs poorly and is generally considered unreliable.

Autocorrelation is an approach to overcoming the clutter problem. Typically, a window of data is selected and convolved with subsequent data to form correlation estimates. The lag which yields the largest correlation estimate above some threshold is chosen as the period of the heart cycle. This technique is more reliable than peak detection, although it has several limitations. Problems in detection are caused when the signal is corrupted by somewhat periodic background components. The most common source of these is heavy breathing. It is difficult to filter out breathing because the exact breathing rate (and hence its frequency components) are unknown beforehand and the breathing components are usually located very close to the heart components in the frequency spectrum, (often within a fraction of a Hertz). Further problems are caused by the nature of the heart beat itself. Autocorrelation detection requires periodic signals, which can only loosely describe heart signals. The time interval between beats is rarely constant, so that the length of the data window to be convolved should contain only one or two heart beats. This is in conflict with decreasing the estimate variance by increasing the window length. Additionally, the signature of the heart beat also changes from beat to beat, so when this window is convolved with subsequent data, the differences in signatures often prevent the correlation from exhibiting detectable peaks. Also, autocorrelation detection has flaws when used in a medical instrument. The heart rate estimates it produces result from averages (in a sense) of several heart beat periods. It is therefore difficult to detect erratic heart rates or to rapidly identify trends in the data, such as increasing or decreasing heart rates. Problems in implementation are caused by the large amount of memory and processing required to estimate the heart rate by autocorrelation. Additionally,there is also a significant time delay in producing estimates which is incurred by the need to sample several seconds of data and then compute the autocorrelation.

The fundamental problems in this application have been shown to be manageable. It has been demonstrated that microwaves can be used to detect heart motion and that it is possible to perform detection using a portable instrument [1-3]. Before a final version of the instrument is available, several aspects of the existing instruments need to be improved. Several of the improvements which need to be incorporated are

1) a decrease in processing time
2) a reduction in hardware complexity
3) an improvement in heart rate measurement accuracy

In effecting these improvements, adaptive filtering has been of some use. In particular, adaptive filters have been shown to improve both peak detection and autocorrelation performance. Additionally, heart beat detectors constructed from adaptive filters have been implemented which yield significant improvements in accuracy over autocorrelation while requiring significantly less hardware and processing time.

This presentation will explain the applicability of adaptive filtering to the problem of microwave heart rate measurements. The derivations of several adaptive filters will be presented. A discussion of results will be presented which will compare the behavior of several adaptive algorithms and measure their performance using EKG s as a reference. Finally, possibilities for further investigation will be presented.

INVERSE FILTERING FOR HEART BEAT DETECTION

Adaptive filtering is closely related to system identification and parameter estimation. To provide a framework for the application of adaptive filters to this problem, a potential model for the signal will be given and then it will be shown how adaptive filters can be applied to the model.

The model used is similar to that of an autoregressive process. The assumption behind this model is that the signal can be approximated by the output of an all-pole filter excited by impulses and Gaussian noise (Fig. 1).

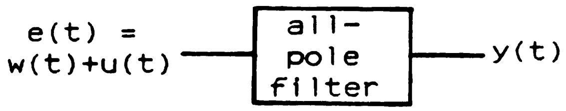

Figure 1. Model of the reflected microwave signal after low pass filtering and sampling

A similar approach is used successfully in speech synthesis and compression; indeed, the similarity of the heart signal to voiced speech motivated this investigation [4].

The all-pole filter in the model will be such that its impulse response matches one heart beat. Obviously, if the input to the filter is an impulse, the filter output will resemble a heart beat. In this way, by driving the filter with an impulse train it is possible to produce an approximation of the heart signal.

To produce aimplitude variations from beat to beat, the impulse train will consist of pulses of varying amplitude. White noise will also be added to the input to introduce random variations in the signal.

The following equation describes this model

$$y(t) = c1\ y(t-1) + c2\ y(t-2) + \ldots + cp\ y(t-p) + e(t) \qquad (1)$$

where $y(t)$ is the output corresponding to the measured data at time t, and ci is the ith coefficient of the order p all-pole filter polynomial. The filter excitation is $e(t) = u(t) + w(t)$. The white Gaussian noise, $w(t)$, has zero mean and variance Rw^2. The impulse train is represented by $u(t)$.

In this model, the filter excitation determines the heart rate. By varying the time between impulses, a time varying heart rate is produced. As such, the occurences of non-zero values of $u(t)$ will correspond to the start of a heart beat. A possible formula for this could be $u(t)=h(t)\ d(t)$. The process $d(t)$ would take on value 1 if a heart beat has occurred since the last data sample was taken. Otherwise it is zero. A way to generate this process could be to take the first difference of a counting process which corresponds to the number of heart beats which have occurred up to the present time. For instance, if $x(t)$ equals the number of heart beats which have occurred in the time interval $[0,t]$, $d(t) = x(t) - x(t-1)$. The time varying coefficient $h(t)$ will allow the driving process to have time varying amplitude.

According to this model, $u(t)$ is non-zero only at the start of a heart beat. As such it contains the information necessary to extract the heart rate from the signal.

The extraction of the pulse train $u(t)$ is an inverse filtering or deconvolution problem which can be viewed in the following way

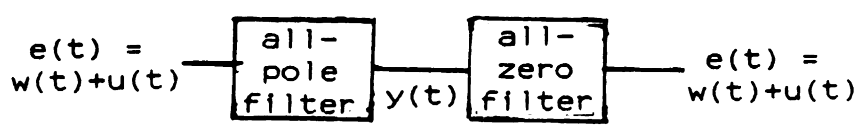

Figure 2. Use of Inverse Filtering to Find the Model Excitaton

Taking the Z-transform of (1),

$$Y(z) \times C(z^{-1}) = U(z) + W(z)$$

$$Y(z) = [C(z^{-1})]^{-1} \times [U(z) + W(z)] \ .$$

where $C(z^{-1}) = 1 - c1\ z^{-1} - \ldots - cp\ z^{-p}$.

If the order of the filter polynomial $C(z)$ and its coefficients were known prior to implementation, inverse filtering would be trivial. However, the filter coefficients of $C(z)$ are unknown and must be determined in real time. The order of $C(z)$ has been found to be sufficiently general from subject to subject to be determined prior to implementation.

There are a number of methods available for determining the coefficients of the inverse filter. The approach described in this paper was motivated by the use of adaptive FIR filters as whitening filters for autoregressive (AR) processes. For AR processes, parameter identification is closely related to the theory of linear prediction [5]. If, in this model, the filter excitation e(t) were white, y(t) would be an AR process. Regression on past data samples would then yield white prediction errors. In this sense, a filter whose output is the least squares prediction error is a whitening or inverse filter. Therefore it is necessary to find a way to do least squares prediction without detailed a priori knowledge of the signal.

PREDICTION

The one-step ahead predictor is given by the equation

$$y\hat{}(t|t-1,\ldots,t-p) = -a(1,p)\ y(t-1)-\ldots-a(p,p)\ y(t-p) \qquad (2)$$

This equation describes a p-th order, one-step ahead predictor. The predictor coefficients are defined with the negative sign so that the prediction error can be defined

$$\begin{aligned} e(p,t) &= y(t) - y\hat{}(t|t-1,\ldots,t-p) \\ &= y(t) + a(1,p)y(t-1) + \ldots + a(p,p)y(t-p) \qquad (3) \end{aligned}$$

This equation describes the p-th order prediction error at time t. The error e(p,t) is the result of convolution of the p data points and the p predictor coefficients. This operation can be realized by a finite impulse response (FIR) filter which will be called the prediction error filter.

Some criterion for the minimization of the prediction errors must be imposed on this filter. A set of coefficients $\{ai, i=1,p\}$ which satisfy this criterion will be considered an optimum linear predictor. The error criterion which will be imposed is the minimization of the mean squared error, J,

$$J = E\{e(p,t)^2\} \qquad (4)$$

where E is the expected value operator.

A criterion equivalent to minimizing J is the orthogonality principle [7].

$$E\{e(p,t)y(t-i)\} = 0 \qquad i=1,p \qquad (5)$$

By demanding that the orthogonality principle be satisfied, J is minimized. This can be shown by differentiating (4) with respect to its coefficients and setting the result equal to zero

$$d/d(a(i,p))\ (E\{e(p,t)\}) =$$

$$d/d(a(i,p))\ E\{(y(t)+a(1,p)\ y(t-1)+...+a(p,p)\ y(t-p))^2\} =$$

$$E\{2(a(1,p)\ y(t-1)+...+a(p,p)\ y(t-p))y(t-i)\} = \quad 0$$

$$E\{\ e(p,t)y(t-i)\ \} = 0 \qquad i=1,p \tag{6}$$

Using the orthogonality condition a set of equations can be derived which can be solved to find the optimum predictor coefficients. This set of equations is called the predictor normal equations [27]. Multiplying (3) by $y(t-j)$ for $j = 1,...,p$ yields

$$y(t)y(t-j) + a(1,p)\ y(t-1)y(t-j) +...+$$
$$a(p,p)\ y(t-p)y(t-j) = e(p,t)y(t-j) \tag{7}$$

Taking the expected value of this equation and using the orthogonality principle yields

$$E\{y(t)y(t-j)\}+a(1,p)E\{y(t-1)y(t-j)+...+a(p,p)E\{y(t-p)y(t-j)\}$$
$$= 0 \tag{8}$$

Defining $R(i-j) = E\{y(t-i)y(t-j)\}$, this last equation can be written

$$R(j) + a(1,p)\ R(1-j) + \ ...\ + a(p,p)\ R(p-j) = 0 \tag{9}$$

(since the process $y(t)$ is assumed to be real, $R(j) = R(-j)$).

For $j = 1,...,p$ equation (6) can be expressed in matrix form

$$\begin{bmatrix} R(1) & R(0) & R(1) & \dots & R(p-1) \\ R(2) & R(1) & R(0) & \dots & R(p-2) \\ : & & & & : \\ : & & & & : \\ R(p) & R(p-1) & & \dots & R(0) \end{bmatrix} \begin{bmatrix} 1 \\ a(1,p) \\ : \\ : \\ a(p,p) \end{bmatrix} = 0 \qquad (10)$$

This system of equations can be increased to p+1 equations by adding in the equation for the variance of the prediction error. Doing so will simplify the structure of the matrix on the left and allow for easier solution of these equations. Squaring equation (3) and taking expected values yields the variance of the error e(p,t) which will be defined

$$Re(p) = E\{e(p,t)^2\}$$

$$\begin{aligned} Re(p) = E\{&(y(t) + a(1,p)\, y(t-1) + \dots + a(p,p)\, y(t-p)) \\ &(y(t) + a(1,p)\, y(t-1) + \dots + a(p,p)\, y(t-p))\} \\ = E\{&y(t)(y(t) + a(1,p)\, y(t-1) + \dots + a(p,p)\, y(t-p)) + \\ &a(1,p)\, y(t-1)(y(t) + a(1,p)\, y(t-1) + \dots + a(p,p)\, y(t-p)) + \\ &: \\ &a(p,p)\, y(t-p)(y(t) + a(1,p)\, y(t-1) + \dots + a(p,p)\, y(t-p))\} \\ = {}&R(0) + a(1,p)\, R(1) + \dots + a(p,p)\, R(p) + \\ &E\{a(1,p)\, y(t-1)e(p,t) + \dots + a(p,p)\, y(t-p)e(p,t)\} \qquad (11) \end{aligned}$$

By the orthogonality principle, the last term equals zero.

$$Re(p) = R(0) + a(1,p)\, R(1) + \dots + a(p,p)\, R(p) \qquad (12)$$

Equation (12) can be included in the set of equations given by (10) so that they take the form

$$\begin{bmatrix} R(0) & R(1) & \cdots & R(p) \\ R(1) & R(0) & & R(p-1) \\ \vdots & & & \\ \vdots & & & \\ R(p) & R(p-1) & \cdots & R(0) \end{bmatrix} \begin{bmatrix} 1 \\ a(1,p) \\ \\ \\ a(p,p) \end{bmatrix} = \begin{bmatrix} Re(p) \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix} \qquad (13)$$

Using matrix notation, equation (9) can be written compactly as Rp A = Q where the Rp is the covariance matrix and Q is the vector on the right hand side.

When these equations are written in the form given by (10), they are called the Yule-Walker equations [9,10]. When the additional error variance equation is included, the form (13) is sometimes called the augmented Yule-Walker equations.

Solution techniques for these equations make use of the symmetric structure of the matrices to avoid performing standard matrix inversion. The matrix equation (10) which has p scalar equations in p unknowns can be solved using Durbin's recursive procedure. This method requires 2p memory locations and p^2 operations as opposed to general matrix inversion which requires on the order of p^3 operations [26]. This method is useful for off-line analysis of data in parameter estimation and order determination routines [11,12,13].

ADAPTIVE FILTERING

In the previous presentation of methods for obtaining least mean squares predictors, the second order statistics were assumed to be known. The autocorrelations up to a lag of p were needed to form the Yule-Walker equations for the p-th order predictor.

For some processes, these statistics may be known beforehand. If so, it is possible to design a predictor beforehand. Often, however, detailed knowledge about a process is not available before actual observations are made. In these cases, estimates of the autocorrelation can be made from the sampled data. The procedure of estimating the statistics of a process and forming a predictor from the estimates will be called adaptive prediction. The related structure which produces the prediction errors will be called the adaptive filter.

Adaptive prediction is also useful for some types of non-stationary data. If the process statistics are slowly time varying, an adaptive filter may be able to track them. A predictor can be constructed, although it will not be strictly correct for non-stationary data.

An objective of the adaptive filter will be short processing time. In some situations, this is not a crucial consideration. For example, if new data arrives infrequently, say monthly, efficiency in estimating process statistics and solving the Yule-Walker equations is not important. In the biomedical applications addressed here, medium speed algorithms are most suitable. Real time analysis and restrictions on hardware complexity rule out time consuming techniques and techniques which require storage of large amounts of data. However, the sampling rates are low enough (~100 Hz) that exactness need not be sacrificed to speed. This means that the algorithms used will be "exact solution" techniques which produce exact solutions to the matrix inversion problem given the statistical estimates. The alternative would have been to use less complex ,gradient descent algorithms which produce estimates which eventually converge to the desired solution [14,15]. The advantage is that the exact solution algorithms have better tracking and start up behavior over the gradient algorithms. This will be advantageous when processing non-stationary data and in shortening the time required to produce estimates.

To summarize, the adaptive predictor objective is to estimate the statistics of the process, solve the predictor equations, and produce a prediction. This will be done each time a new sample becomes available and should be completed before the next sample appears.

Since the statistics of the process are not known, the process to be predicted will be assumed to be ergodic, and expectation operations in the previous stochastic derivation will be replaced by time averages. The error criterion to be minimized changes from E{e(t)^2} to

$$J = e(p,0)^2 + \ldots + e(p,t)^2 \tag{14}$$

a time average approximation where the multiplicative constant is dropped. The following derivations follow the presentation given in [16].

Using matrix notation, the prediction error equation (1) can be written

$$\begin{bmatrix} e(p,0) \\ e(p,1) \\ : \\ : \\ : \\ : \\ e(p,t) \end{bmatrix} = \begin{bmatrix} y(0) \\ y(1) \\ : \\ : \\ : \\ : \\ y(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 & \ldots & 0 \\ y(0) & 0 & \ldots & 0 \\ y(1) & y(0) & 0 \ldots & 0 \\ : & & & : \\ : & & & y(0) \\ : & & & : \\ y(t-1) & & \ldots & y(t-p) \end{bmatrix} \begin{bmatrix} a(1,p) \\ a(2,p) \\ : \\ : \\ : \\ : \\ a(p,p) \end{bmatrix} \tag{15}$$

This arrangement of the (txp) data matrix is called the pre-windowed form [6]. For simplicity, the data matrix will be denoted as Yp(t).

The error criterion can be expressed in matrix form

$$e(p,0)^2 + \ldots + e(p,t)^2 = [e(p,0) \ \ldots \ e(p,t)] \begin{bmatrix} e(p,0) \\ : \\ e(p,t) \end{bmatrix} \tag{16}$$

Rewriting (15) in the minimum prediciton error form, i.e. setting the prediction errors to zero, the following equation can be obtained, where ′ denotes transposition,

$$-Yp(t)' \begin{bmatrix} y(0) \\ : \\ y(t) \end{bmatrix} = Yp(t)' \ Yp(t) \begin{bmatrix} a(1,p) \\ : \\ a(p,p) \end{bmatrix} \tag{17}$$

The solution for the coefficient vector is given by using the generalized matrix inverse

$$-(Yp(t)'\ Yp(t))^{-1}\ Yp(t)' \begin{bmatrix} y(0) \\ y(1) \\ : \\ y(t) \end{bmatrix} = \begin{bmatrix} a(1,p) \\ a(2,p) \\ : \\ a(p,p) \end{bmatrix} \qquad (18)$$

The notation will be simplified further by expressing the data vector in this equation as y(0:t). Equation (18) is the solution for linear regression on past data of the process y(t). Therefore, standard recursive techniques such as Gaussian least squares are applicable. These regression techniques solve the general estimation problem of minimizing the error criterion (14).

The matrix equation (18) has the term (Yp(t)' Yp(t)) which can be expanded into

$$Yp(t)'\ Yp(t) =$$

$$\begin{bmatrix} y(0)^2 + \ldots + y(t-1)^2 & \ldots & y(0)y(p-1) + \ldots + y(t-1)y(t-p) \\ y(0)y(1) + \ldots + y(t-2)y(t-1) & \ldots & y(0)y(p-2) + \ldots + y(t-2)y(t-p) \\ : & & : \\ : & & : \\ : & & : \\ y(0)y(p-1) + \ldots + y(t-p)y(t-1) & \ldots & y(0)^2 + \ldots + y(t-p)^2 \end{bmatrix} \qquad (19)$$

This matrix is called the sample covariance matrix for pre-windowed data, noting that the processes being filtered here are all zero mean because a low frequency (~0.5 Hz) high pass filter is used prior to sampling.

SIMPLE FIRST ORDER PREDICTOR

The first order predictor is simple to implement directly from the Yule-Walker equations [17]. From equation (13) (setting p = 1),

$$y\hat{}(n) = -a\, y(n-1)$$

$$a = -\, Ry(1)/Ry(0) \qquad (20)$$

$$Re(1)^2 = (1 - a^2)\, Ry(0)$$

where Ry(1) will be approximated by

$$Ry(1) = [q^t\, y(0)y(1) + \ldots + q\, y(t-1)y(t-2) + y(t)y(t-1)]/t$$

We will implement a recursive approximation using the time update equation for Ry(i) at time t

$$Ry(i,t) = q^*\, Ry(i,t-1) + x(t-i)^*x(t)/t \qquad (21)$$

The forgetting factor q introduces an exponential weighting to the data which decreases the influence of data further in the past. It is assumed that the signal is short term stationary in the middle of the heart beat, so q will be chosen such that non-stationary effects of the previous heart beat will be discarded before the next beat appears. To chose q, the measure of the effective windowing length, $n = q/(1-q)$ is used [18]. For a length of 0.5 second, $q = 0.98$ for a sampling rate of about 100 Hz. This value of n will be substituted for t in equation (21).

The equation (21) has another interpretation. It can be considered as a low pass filter, where q locates the cut off frequency, $w = 1-q$ [25], and w is normalized by the sampling frequency. A method of choosing q would then be to locate the cut off frequency above the major components in the power spectrum. Note that these comments about q apply to exponential windowing in general.

LATTICE FILTERS

Another technique for solving the prediction problem is Levinson's algorithm which is a recursive solution for the matrix equation (13). A full description of this technique starting with the case of known second order statistics and proceeding to the adaptive case will be presented because this approach forms the basis of a useful form of adaptive filter which will be introduced later. This approach introduces a set of backward prediction errors

$$r(p,t-1) = y(t-p-1) - \hat{y}(t-p-1 \mid y(t-p),\ldots,y(t-1)) \qquad (22)$$

where the backward prediction is defined

$$\hat{y}(t-p-1 \mid y(t-p),\ldots,y(t-1)) = -b(1,p)\, y(t-p) - \ldots - b(p,p)\, y(t-1) \qquad (23)$$

The backward prediction error can be found using a prediction error filter similar to that of the forward prediction error. Combining (22) and (23),

$$r(p,t-1) = y(t-p-1) + b(1,p)\, y(t-p) + \ldots + b(p,p)\, y(t-1) \qquad (24)$$

Both the forward and backward prediction predictors use the same block of data, the values of $y(t)$ from $t-p$ to $t-1$.

In the same way as the augmented Yule-Walker equations (13) were derived for the forward predictor, similar results can be obtained for the backward predictor. These equations take the form

$$\begin{bmatrix} R(0) & R(1) & \cdots & R(p) \\ R(1) & R(0) & & R(p-1) \\ \vdots & & & \\ \vdots & & & \\ R(p) & R(p-1) & \cdots & R(0) \end{bmatrix} \begin{bmatrix} b(p,p) \\ \\ \\ b(1,p) \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ Rr(p) \end{bmatrix} \qquad (25)$$

where $Rr(p)$ is the variance of the p-th order backward prediction error

$$Rr(p) = E\{r(p,t)^2\} \qquad (26)$$

Levinson's algorithm solves equations (13) and (25) simultaneously and recursively. The algorithm is recursive in that, if given the optimum p-th order predictors, the algorithm produces the predictor of order p+1. The derivation for the recursions is presented here, as outlined in [15].

Assuming that the forward and backward p-th order predictors are known, they must satisfy equations (13) and (25). These two equations can be combined and expressed in the following form

$$\begin{bmatrix} R(0) & R(1) \cdots R(p) & R(p+1) \\ R(1) & R(0) & R(p) \\ \vdots & & \vdots \\ \vdots & & \vdots \\ \vdots & & \vdots \\ R(p+1) & \cdots R(1) & R(0) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ a(1,p) & b(p,p) \\ \vdots & \vdots \\ \vdots & \vdots \\ a(p,p) & b(1,p) \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} Re(p) & Lr(p+1) \\ 0 & 0 \\ \vdots & \vdots \\ \vdots & \vdots \\ 0 & 0 \\ Le(p+1) & Rr(p) \end{bmatrix} \qquad (27)$$

The terms $Lr(p+1)$ and $Le(p+1)$ arise from expanding the (p+1 x p+1) matrix equations (13) and (25) into a set of (p+2 x p+2) matrix equations.

If an operation can be found which, when performed on (14), forces Le(p+1) and Lr(p+1) to zero, the transformed set of equations will then satisfy equations (13) and (25) for the predictor of order p+1. The desired set of equations has the form

$$\begin{bmatrix} R(0) & \ldots R(p) & R(p+1) \\ R(1) & & : \\ : & & : \\ : & & : \\ R(p+1) \ldots & R(1) & R(0) \end{bmatrix} \begin{bmatrix} 1 & b(p+1,p+1) \\ a(1,p+1) & b(p,p+1) \\ : & : \\ : & : \\ a(p+1,p+1) & 1 \end{bmatrix} = \begin{bmatrix} Re(p+1) & 0 \\ 0 & 0 \\ : & : \\ : & : \\ 0 & Rr(p+1) \end{bmatrix} \quad (28)$$

The operation which transforms (27) into the desired form (28) is post-multiplication by

$$\begin{bmatrix} 1 & -Lr(p+1)/Re(p) \\ -Le(p+1)/Rr(p) & 1 \end{bmatrix} = \begin{bmatrix} 1 & -ke(p+1) \\ -kr(p+1) & 1 \end{bmatrix} \quad (29)$$

where the forward and backward reflection coefficients, kr(p+1) and ke(p+1), are defined as ke(p+1) = -Lr(p+1)/Re(p) and kr(p+1) = Le(p+1)/Rr(p) . Writing this operation explicitly,

$$\begin{bmatrix} R(0)\ldots & R(p+1) \\ : & : \\ : & : \\ : & : \\ R(p+1)\ldots & R(0) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ a(1,p) & b(p,p) \\ : & : \\ a(p,p) & : \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -ke(p+1) \\ -kr(p+1) & 1 \end{bmatrix} = \begin{bmatrix} Re(p+1) & 0 \\ 0 & : \\ : & : \\ : & 0 \\ 0 & Rr(p+1) \end{bmatrix} \quad (30)$$

Equating (30) to (28) yields

$$\begin{bmatrix} R(0) & & R(p+1) \\ : & & : \\ : & & : \\ : & & : \\ R(p+1) & & R(0) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ a(1,p) & b(p,p) \\ : & : \\ a(p,p) & : \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -ke(p+1) \\ -kr(p+1) & 1 \end{bmatrix} =$$

$$\begin{bmatrix} R(0) & R(1) & \dots & R(p) & R(p+1) \\ R(1) & R(0) & & & \\ : & & & & \\ : & & & & \\ R(p+1) & R(p) & \dots & R(1) & R(0) \end{bmatrix} \begin{bmatrix} 1 & b(p+1,p+1) \\ a(1,p+1) & b(p,p+1) \\ : & : \\ : & : \\ a(p+1,p+1) & 1 \end{bmatrix} \quad (31)$$

Pre-multiplication by the inverse of the covariance matrix yields the order update recusions

$$\begin{bmatrix} 1 & 0 \\ a(1,p) & b(p,p) \\ : & : \\ a(p,p) & : \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -ke(p+1) \\ -kr(p+1) & 1 \end{bmatrix} = \begin{bmatrix} 1 & b(p+1,p+1) \\ a(1,p+1) & b(p,p+1) \\ : & : \\ : & : \\ a(p+1,p+1) & 1 \end{bmatrix} \quad (32)$$

The Levinson order update recursions given in equation (32) can be expressed directly in a digital filter in a form called a lattice filter. To see how equation (32) can be incorporated into a filter, take the Z-transform of the prediction error equations (3) and (24)

$$e(p,z) = Ap(z)\ Y(z) \quad (33)$$

$$r(p,z) = z\hat{}-1\ Bp(z)\ Y(z)$$

where $Ap(z) = 1 + a(1,p)\ z\hat{}-1 + \dots + a(p,p)\ z\hat{}-p$

and $Bp(z) = b(p,p)\ z\hat{}-1 + \dots + b(p,1)\ z\hat{}-(p-1) + z\hat{}-p$

Premultiplying (33) by [1,z^-1,...,z^-p,z^-(p+1)], the order update recursion equation (32) can be written in Z-transform notation

$$[1,z^{-1},\dots,z^{-p},z^{-(p+1)}]\begin{bmatrix} 1 & 0 \\ a(1,p) & b(p,p) \\ \vdots & \vdots \\ a(p,p) & 1 \end{bmatrix}\begin{bmatrix} 1 & -ke(p+1) \\ -kr(p+1) & 1 \end{bmatrix} =$$

$$[1,z^{-1},\dots,z^{-p},z^{-(p+1)}]\begin{bmatrix} 1 & b(p+1,p+1) \\ a(1,p+1) & b(p,p+1) \\ \vdots & \vdots \\ a(p+1,p+1) & 1 \end{bmatrix} \qquad (34)$$

$$[Ap(z)\ \ Bp(z)]\begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix}\begin{bmatrix} 1 & -ke(p+1) \\ -kr(p+1) & 1 \end{bmatrix} = [Ap+1(z)\ \ Bp+1(z)] \qquad (35)$$

Premultiplying (35) by Y(z),

$$Y(z)[Ap(z)\ \ Bp(z)]\begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix}\begin{bmatrix} 1 & -ke(p+1) \\ -kr(p+1) & 1 \end{bmatrix} = Y(z)[Ap+1(z)\ \ Bp+1(z)]$$

$$[e(p,z)\ \ r(p,z)]\begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix}\begin{bmatrix} 1 & -ke(p+1) \\ -kr(p+1) & 1 \end{bmatrix} = [e(p+1,z)\ \ \ r(p+1,z)] \qquad (36)$$

Equation (36) describes how the lattice filter is to be implemented. The equation describes the propagation of the prediction errors of predictors of increasing order. Starting with order 0 (p=0), A0 = B0 = 1. The reason for this is that a zero order predictor makes no prediction (the prediction equals zero) and the prediction error is the value to be predicted.

The first stage of the lattice (from order 0 to order 1) is given by

$$[Y(z)\ \ Y(z)]\begin{bmatrix} 1 & 0 \\ 0 & z\hat{}-1 \end{bmatrix}\begin{bmatrix} 1 & -ke(1) \\ -kr(1) & 1 \end{bmatrix} = [e(1,z)\ \ r(1,z)] \qquad (37)$$

The flow diagram for this is

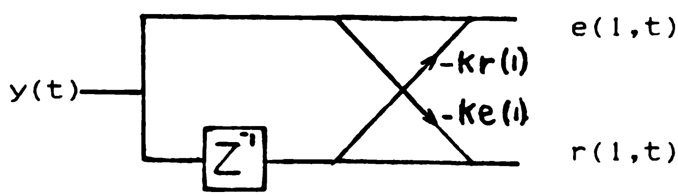

Figure 3. Stage One of the Lattice Filter

The order update from order p = 1 to order p = 2 is given by

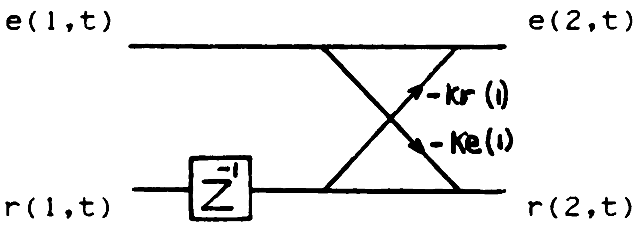

Figure 4a. Stage two of the Lattice Filter

The general lattice of order p is

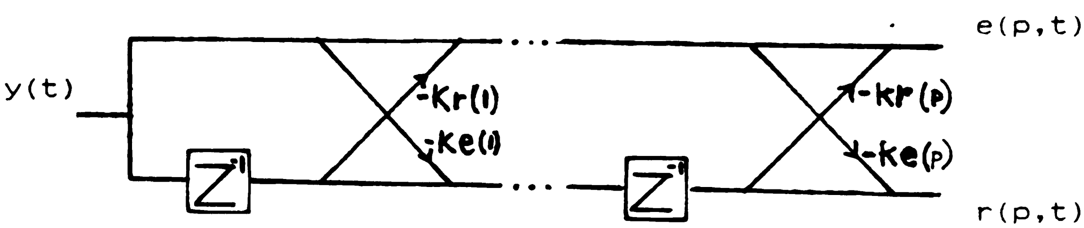

Figure 4b. Lattice Filter

Some inspection of the filter parameters will simplify these realizations. The quantities Lr(p+1) and Le(p+1) are equal. To see this, extract Lr(p+1) and Le(p+1), as defined in (29) from equation (27) using the operations on the matrix Rp+1

$$[1\ a(1,p)\dots a(p,p)\ 0]R_{p+1}\begin{bmatrix}0\\ b(p,p)\\ \vdots\\ b(1,p)\\ 1\end{bmatrix} = [1\ a(1,p)\dots a(p,p)\ 0]\begin{bmatrix}Lr(p+1)\\ 0\\ \vdots\\ \vdots\\ Rr(p+1)\end{bmatrix}$$

$$= Lr(p+1) \qquad (38)$$

and

$$[0\ b(p,p)\dots b(1,p)\ 1]R_{p+1}\begin{bmatrix}1\\ a(1,p)\\ \vdots\\ a(p,p)\\ 0\end{bmatrix} = [0\ b(p,p)\dots b(1,p)\ 1]\begin{bmatrix}Re(p+1)\\ 0\\ \vdots\\ \vdots\\ Le(p+1)\end{bmatrix}$$

$$= Le(p+1) \qquad (39)$$

Since the left hand sides of equations (38) and (39) are transpositions of each other, their scalar products are equal. The notation r and e will be dropped from Lr(p+1) and Le(p+1) and they will be denoted by a single variable L(p+1).

The variables associated with this update recursion also have information about the prediction process. By comparing equations (27) and (30), order update recursions can be found for the prediction error variances

$$Re(p+1) = Re(p) - Lr(p+1)\ kr(p+1)$$

$$Rr(p+1) = Rr(p) - Le(p+1)\ ke(p+1) \qquad (40)$$

The variable L(p+1) is the cross correlation of the forward and backward prediction errors at a lag of 1. This can be seen from expanding (38) into the following form

$$L(p+1) =$$

$$E\{[0\ b(p,p)\ldots b(1,p)\ 1]\begin{bmatrix} y(t) \\ y(t-1) \\ : \\ y(t-p-1) \end{bmatrix}[y(t)\ldots y(t-p-1)]\begin{bmatrix} 1 \\ : \\ a(p,p) \\ 0 \end{bmatrix}\}$$

$$L(p+1) = E\{r(p,t-1)\ e(p,t)\} \qquad (41)$$

The previous derivations assumed that the forward and backward predictors were different, which is not correct. Since least mean squares prediction is analogous to a projection onto the space spanned by the data y(t-1),...,y(t-p) [8], the forward and backward predictors are the same. Also, the prediction error variances Re(p) and Rr(p) are equal, and can be found recursively as in equation (28). This means that the forward and backward reflection coefficients, kr(p+1) and ke(p+1), are also equal. The reflection coefficients can also be used to check the invertibility of the filter. An all zero filter can be inverted if and only if all its zeros are inside the unit circle in the Z-plane, i.e. if it is minimum phase. It can be shown [20] that this is true for the least squares predictor if and only if $|k(p+1)| < 1$.

These equalities are not always imposed on the realization in Figure 5. When the filter is implemented adaptively, the parameters are often computed independently of each other to allow for non-stationarities in the data and errors in approximating the parameters.

The actual predictors Ap(z) .and Bp(z) are not immediately available from the lattice filter. There is, however, a simple way to obtain them. Since the lattice filter has the same impulse response as the predictors in equations (3) and (22), the predictor coefficients can be found from the impulse response of the lattice filter. This makes use of the knowledge that the i-th value of an FIR impulse response is the i-th filter coefficient. To obtain the coefficients of forward predictor of order j (j <= p), observe the e(j,t) node in Figure 3. If the states of the lattice (the nodes r(i,t-1)) are set to zero, the response at node e(j,t) to an impulse applied at y(t) at time t = 0 will be 1,a(1,j),...,a(j,j) which are the coefficients of the predictor.

The order p prediction y^(t) can be found by inspection without actually finding the predictor Ap(z), since y^(t) = y(t) - e(p,t), from equation (3).

By implementing a lattice filter of order p, the predictions and prediction errors of all the lower order predictors Ai(z), i=1,p are available by inspection.

NORMALIZED LEAST SQUARES LATTICE FILTER

Several forms of adaptive lattice filters are presented in [16]. One of these which has proven particularly useful is the normalized, pre-windowed lattice filter. As the name suggests, the filter is in the lattice form of Levinson's Recursion (Figure 5). The pre-windowed qualification means that the data used to form the predictor is taken from the time interval [0,t-1]. The filter is called normalized because the prediction errors propagating through it are divided by their standard deviations. Unnormalized lattice forms exist and it is possible to obtain the unnormalized forms from the normalized forms, and vice versa. The normalized form is preferable in that it requires fewer operations at each recursion than the unnormalized filter.

Friedlander's derivation is simpler and more straight forward than the derivation previously presented by Lee, Morf and Friedlander [21]. The latter derivation is based on a Hilbert space approach which yields several useful interpretations not immediately obvious in the derivation used by Friedlander.

Friedlander's derivation proceeds by considering the least squares predictor in equation (18) as a projection operator onto the space spanned by the columns of Yp(t). Then, he presents matrix operations which update the projection operator of the (txp) data matrix to the projection operator of a data matrix increased in order or time. By combining these time and order update operations, a recursive solution for the lattice filter is derived. The details of the derivation are included in the Appendix.

Each parameter in the filter can be computed using recursive formulas. To ease the notation, functions will be defined for the recursions.

$$F(a,b,c) = [1-c^2]^{-1/2} [a - bc][1-b^2]^{-1/2} \quad (42)$$

$$G(a,b,c) = [1-c^2]^{1/2} a[1-b^2]^{1/2} + cb \quad (43)$$

Using this notation the filter recursions can be expressed in the form

$$k(p+1,t) = G(k(p+1,t-1), \tilde{r}(p,t-1), \tilde{e}(p,t)) \quad (44)$$

$$\tilde{e}(p+1,t) = F(\tilde{e}(p,t), \tilde{r}(p,t-1), k(p+1,t))$$

$$\tilde{r}(p+1,t) = F(r(p,t-1), e(p,t), k(p+1,t)$$

The recursions are performed each instant a data value becomes available. The initial conditions are to set all parameters equal to 0 at time $t = -1$.

Lee, Morf and Friedlander's derivation allows incorporation of an exponential weighting factor into the data which reduces the influence of past data values. They show that an exponential window can be introduced by windowing and normalizing the data. instant. That is, the variance of the zero order prediction error (the zero order prediction error is the data value) is given an exponential taper by using the estimate

$$s(t) = q\, s(t-1) + y(t)^2 \qquad 0 < q < 1 \quad (45)$$

Then, e~(0,t) = r~(0,t) = y(t)/sqrt(s(t))

The complete filter algorithm is then (46)

at time t:

obtain y(t)

s(t) = q s(t-1) + y(t)^2

e~(0,t) = r~(0,t) = y(t)/sqrt(s(t))

For m = 0, p-1 :

k(m+1,t) = G(k(m+1,t-1),r~(m,t-1),e~(m,t))

e~(m+1,t) = F(e~(m,t),r~(m,t-1),k(m+1,t))

r~(m+1,t) = F(r(m,t-1),e(m,t),k(m+1,t)

The filter has another useful parameter which does not arise immediately from Friedlander's derivation. This variable propagates unseen in the normalized lattice filter (similarly to the exponential weighting l). This variable is a likelihood variable which detects changes in the statistics of the input process. If the process statistics are changing, the filter parameters are allowed to vary more rapidly to track the process. Friedlander gives a method for computing the likelihood parameter, g(t),

1 - g(t-1) = (1 - r~(0,t-1)^2)(1 - r~(1,t-1)^2)*...*

(1 - r~(p-1,t-1)^2) (47)

for the predictor of order p-1.

Where g(t) is defined

g(t-1) = y(t-p:t-1)'((Rp-1)^-1)y(t-p:t-1) (48)

This is the exponential term in the zero mean, joint Gaussian distribution for p-1 variables and Rp-1 is the covariance matrix as defined in (9).

Pseudo Linear Regression for Least Squares Prediction

A predictor can be implemented by performing matrix inversion using an algorithm based on the Matrix Inversion Lemma [22]. This lemma states that if a matrix can be expressed as

$$P(t-1)^{-1} = P(t-2)^{-1} + h(t-1)\, h(t-1)' \qquad (49)$$

then the inverse of the matrix can be found by

$$P(t-1) = P(t-2) - P(t-2)\, h(t-1)\, h(t-1)'\, P(t-2) / (1 + h(t-1)'\, P(t-2)\, h(t-1)) \qquad (50)$$

This recursion is similar to those found for the adaptive lattice algorithm. Their purpose is to allow matrix inversion by a recursive algorithm which requires only scalar division.

The predictor which will be used is designed to minimize the error criterion [23]

$$J = (e(p,1)^2 + \dots + e(p,t)^2)/t \qquad (51)$$

where $e(t) = y(t) - \hat{y}(t)$.

The estimate $\hat{y}(t)$ is chosen as a linear function of available data.

The data used to predict $y(t)$ is taken from the time interval $[t-1,t-p]$ and the predictor is formed using the time varying coefficients $a1(t),\dots,ap(t)$. The prediction is then given by

$$y^\wedge(t) = [y(t-1),\ldots,y(t-p)] \begin{bmatrix} a1(t) \\ \vdots \\ ap(t) \end{bmatrix} = h(p,t-1)\ A(p,t) \qquad (52)$$

The algorithm used to update the filter coefficients is the least squares form of the Pseudo Linear Regression algorithm [24]. This is essentially the general recursive least squares algorithm for linear regression applied to the parameters of an FIR predictor . The algorithm is

$$A(p,t) = A(p,t-1) + P(t-2)\ h(p,t-1)'\ e(p,t)/\ m(t-1)$$
$$P(t-1) = P(t-2) - P(t-2)\ h(p,t-1)'\ h(t-1)\ P(t-2)/\ m(t-1)$$
$$m(t-1) = 1 + h(p,t-1)P(t-2)h(p,t-1)' \qquad (53)$$

This algorithm has very good initial start up performance, but after several iterations, the matrix P(t) becomes small and the parameter estimates do not change much at each recursion. It is possible to apply an exponential window to the data by introducing a scale factor into the recursion for P(t-1) [25]

$$P(t-1) = [(P(t-2) - P(t-2)\ h(t-1)'\ h(t-1)\ P(t-2)/g(t-1)]/a(t-1)$$
$$g(t-1) = a(t-1) + h(p,t-1)P(t-2)h(p,t-1)' \qquad (54)$$

A possible choice for a(t-1) is $a(t-1) = a\ a(t-2) + (1-a)$ where $a = .99$ and $a(0) = .95$. This will apply an exponential window to the initial data and then the window will become constant. To window all the data, let

$a(t-1) = q \quad 0 < q < 1$

The algorithm then takes the form

$$
\begin{aligned}
y\hat{}(t) &= h(p,t-1)\ A(p,t) \\
A(p,t) &= A(p,t-1) + P(t-2)h(p,t-1)'\ e(p,t)/g(t-1) \\
e(p,t) &= y(t) - y\hat{}(t) \\
h(p,t) &= [y(t-1),\ldots,y(t-p)] \\
P(t-1) &= [(P(t-2) - \\
&\qquad P(t-2)\ h(t-1)'\ h(t-1)\ P(t-2)/g(t-1)]/a(t-1) \\
g(t-1) &= a(t-1) + h(p,t-1)P(t-2)h(p,t-1)'
\end{aligned}
\tag{55}
$$

In addition to being based on recursive matrix inversion, this algorithm is similar to the adaptive lattice algorithm in that it is possible to obtain a log-likelihood function from the recursive process. This is so because the matrix P(t) is the inverse of the estimate of the covariance matrix in the pre-windowed data case [19]. In computing the parameter g(t-1), this matrix is pre- and post- multiplied by the data vector so the variable g(t-1) is of the form

$$
g(t-1) = a(t-1) + [y(t-1),\ldots,y(t-p)]\ R_{p,(t-2)}\hat{}-1 \begin{bmatrix} y(t-1) \\ \vdots \\ y(t-p) \end{bmatrix} \tag{56}
$$

in which the second term is the exponential term in the joint distribution for several zero-mean Gaussian variables.

DETECTION OF HEART BEATS IN FILTERED DATA

The previous discussions and derivations have shown how adaptive filters are applicable to the problem remoted heart rate detection. Given the model of the signal presented in equation (1), the heart rate can be found by detecting the non-zero occurences of u(t) which initiate a heart beat. The information, u(t), is present in the model as part of the model excitation e(t) = u(t) + w(t), where w(t) is white noise. The excitation e(t) can be found by inverse filtering, which is equivalent to producing the one-step prediction errors of appropriate order. Since the model parameters are not known, they will be determined adaptively and used to form a predictor. The prediction errors will be used as estimates of e(t).

Since the impulse train u(t) is found in the presence of additive Gaussian white noise, some processing must be done on the prediction errors to find the best estimate of u(t). This task becomes more difficult as the variance of the white noise w(t) increases.

The simplest way to detect occurrences of u(t) is to perform peak detection on the prediction errors. In cases where the standard deviation of w(t) is small relative to the amplitude of u(t), the occurrences of u(t) should be noticeable. Additionally, the adaptive behavior of the algorithms aid detection. Since least squares estimation algorithms produce biased estimates when the model excitation is correlated, the occurrence of u(t) will cause estimation errors and the algorithms will change their tracking behavior to correct the errors. This will result in large prediction errors. This behavior when u(t) occurs aids detection.

This type of detection is illustrated in Figure 5. The signal presented in Figure 5a was obtained by placing the microwave transceiver directly on the chest of a reclining subject. The subject was rested and breathing normally. Peak detection without adaptive filtering would have been adequate for this signal. The effects of adaptive filtering are shown well by this example. The results of filtering by each of the three algorithms are shown in Figure 5b-d. As expected, there were large prediction errors when a heart beat occurred (indicated by arrows on the time scale, which were taken from simultaneous EKG measurements). Additionally, between beats, the signal was whitened.

An instance where both peak detection and autocorrelation perform poorly is when the subject is breathing heavily. Such a signal was taken from a subject who had just exercised strenuously (Figure 6a). The subject was seated three feet from the transceiver when the data was taken. Large movements in the chest, due to the breathing, cause peak detection to fail. Autocorrelation also performs poorly, as shown in Figure 6b, perhaps because of the periodic breathing components or the erratic heart rate.

The results of applying adaptive filtering to this signal are shown in Figure 6c-e. Peak detection of the prediction errors is not an appropriate method of detecting u(t) in this case. Either the standard deviation of w(t) is much larger than u(t), or the signal components are such that the filter tracks wildly. It is possible to apply autocorrelation to the prediction errors however. The results of this are shown in Figure 6f-h.

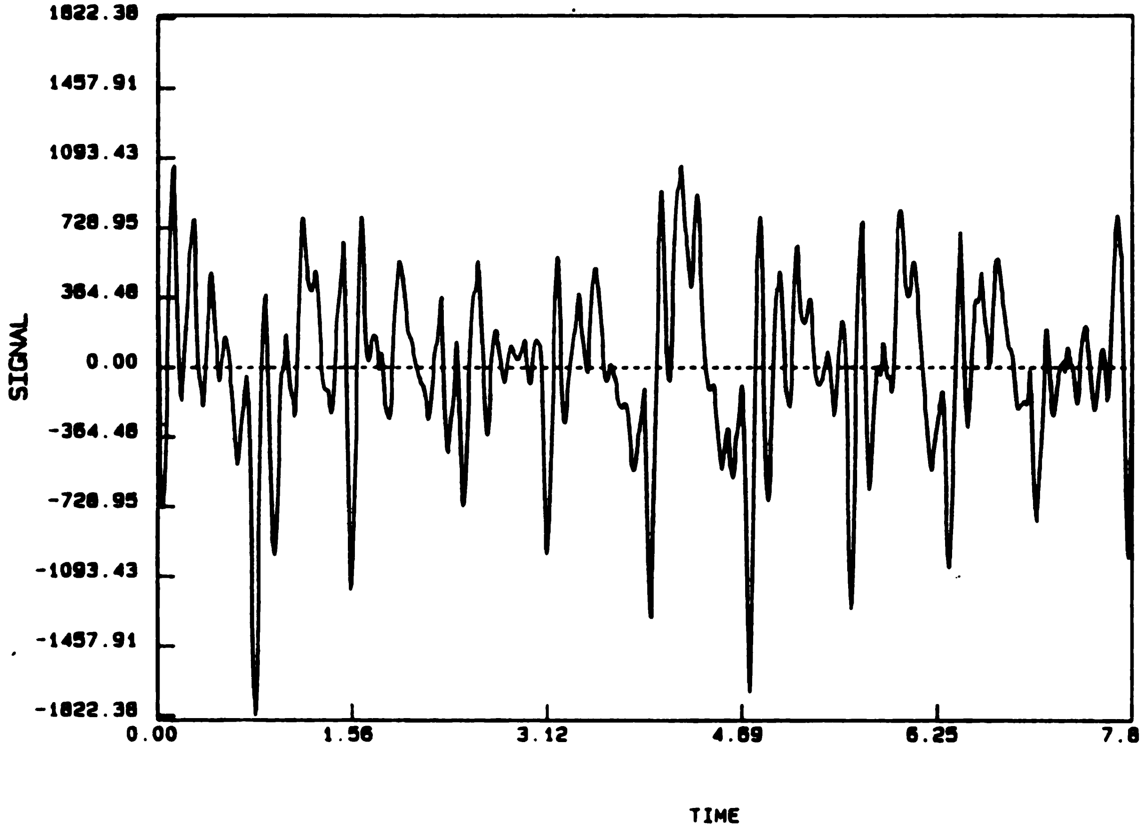

Figure 5a. Microwave Signal Measured by Placing the Transceiver on the Chest of a Resting Subject (sampling rate = 128 Hz)

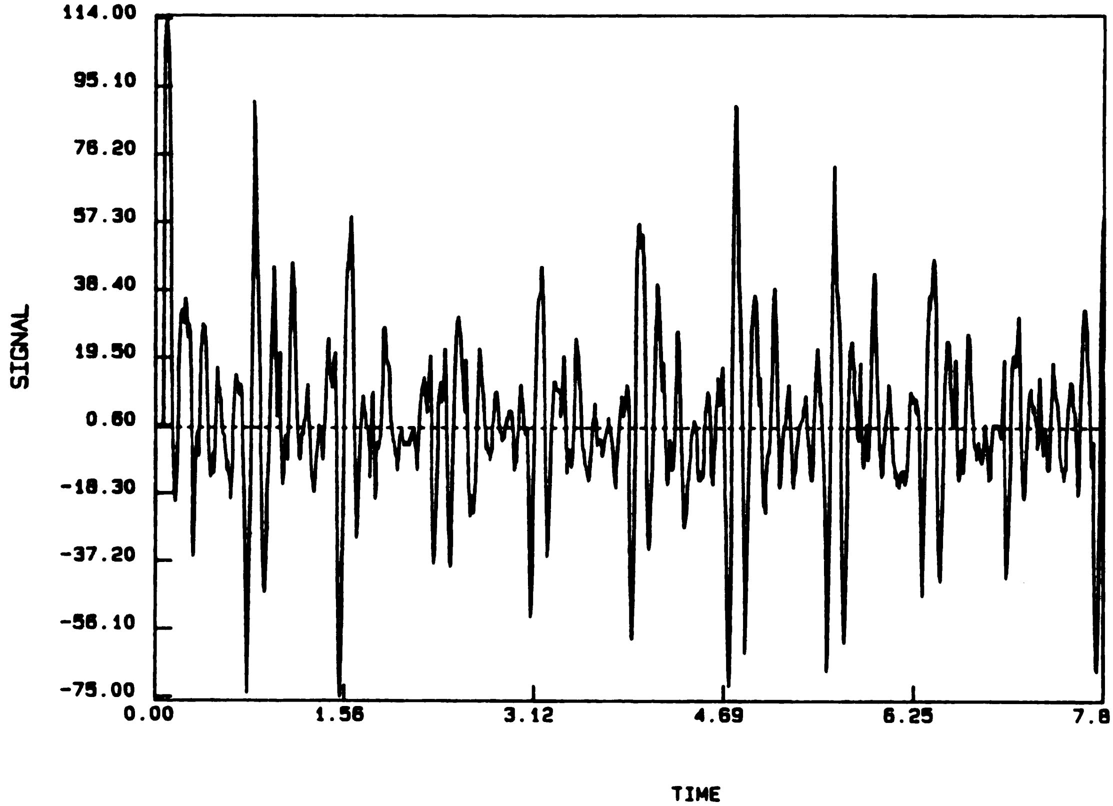

Figure 5b. Prediction Errors of the Simple First Order Predictor, q=.98

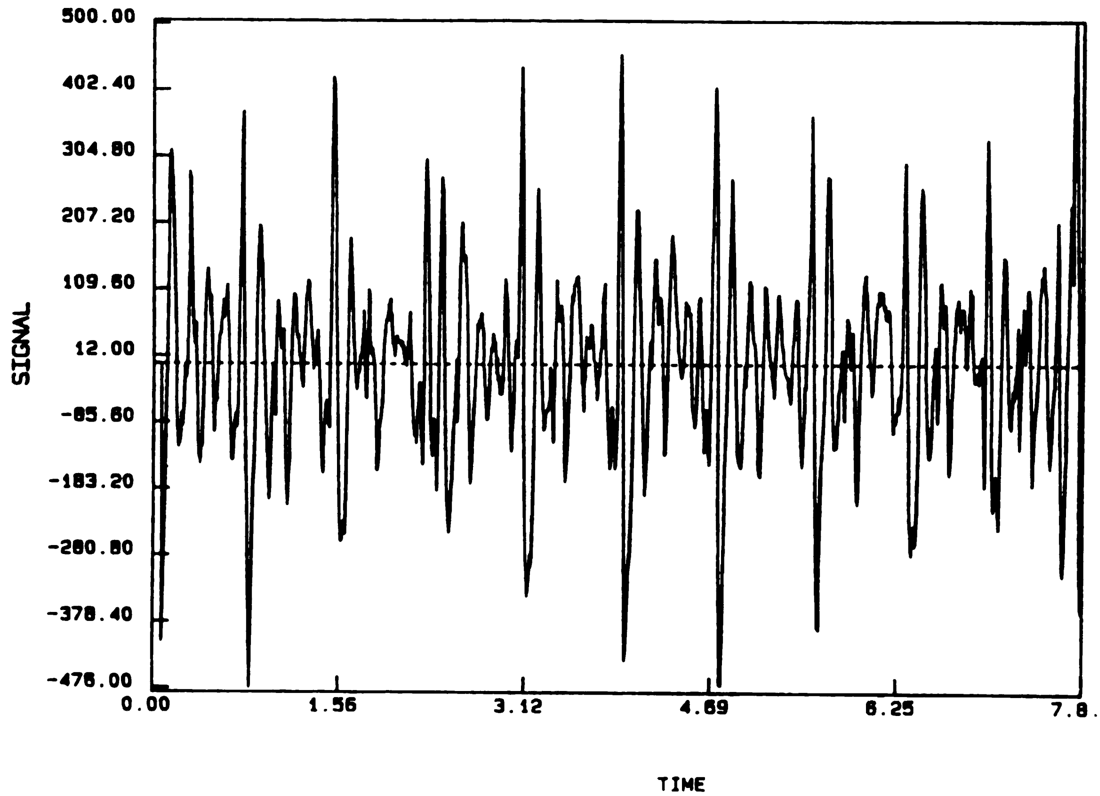

Figure 5c. Prediction Errors of the Normalized Least Squares Lattice Filter (p=1,q=.98)

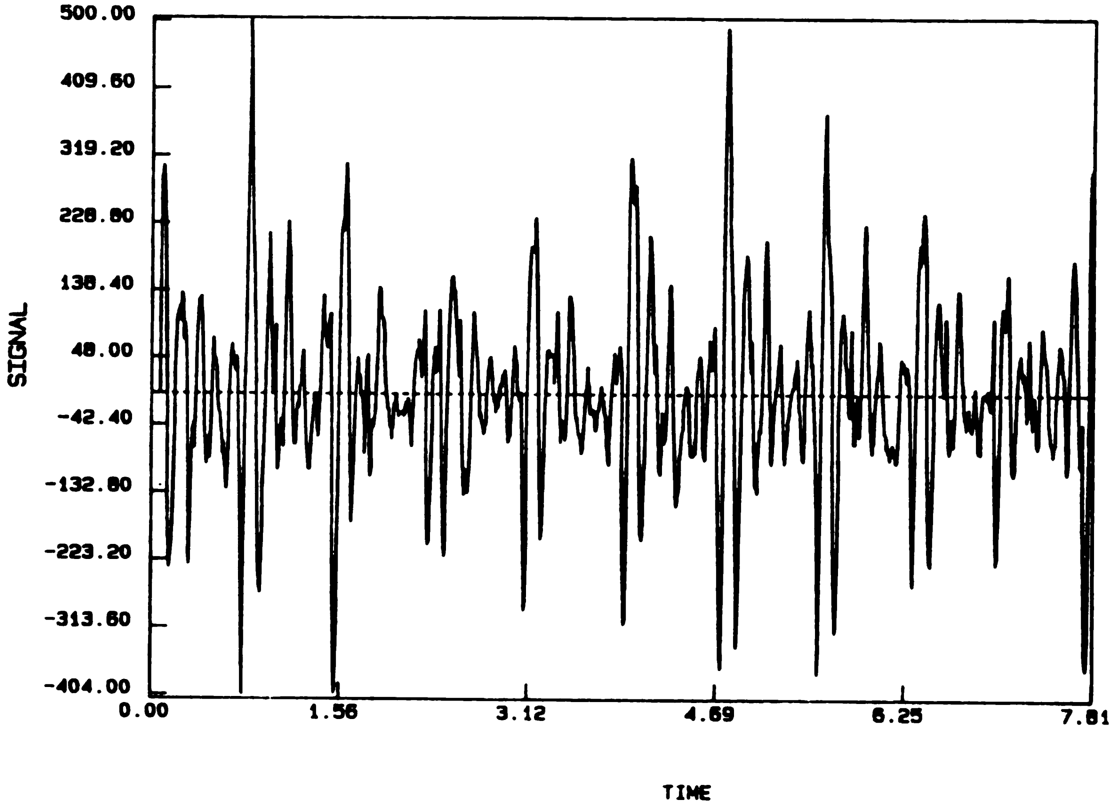

Figure 5d. Prediction Errors of the Pseudo Linear Regression Prediction Algorithm (p=1,q=.98)

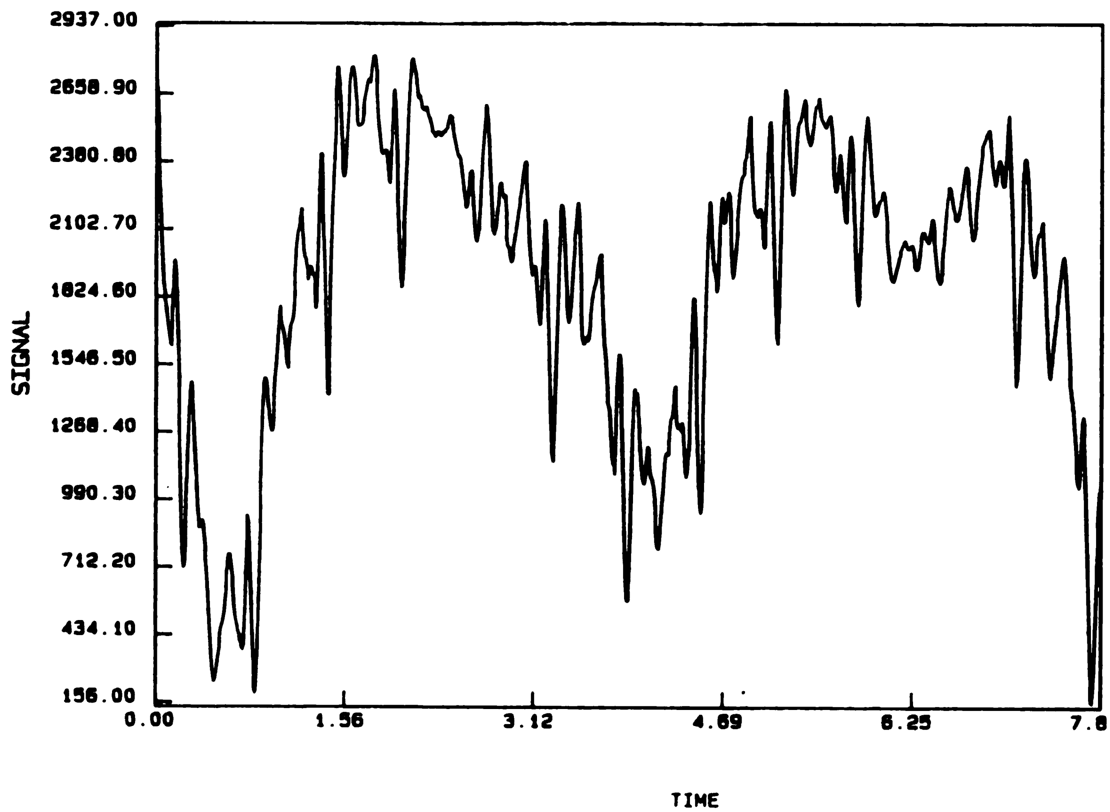

Figure 6a. Microwave Signal Taken from an Exercised Subject Seated Three Feet from the Transceiver
(sampling rate = 128 Hz)

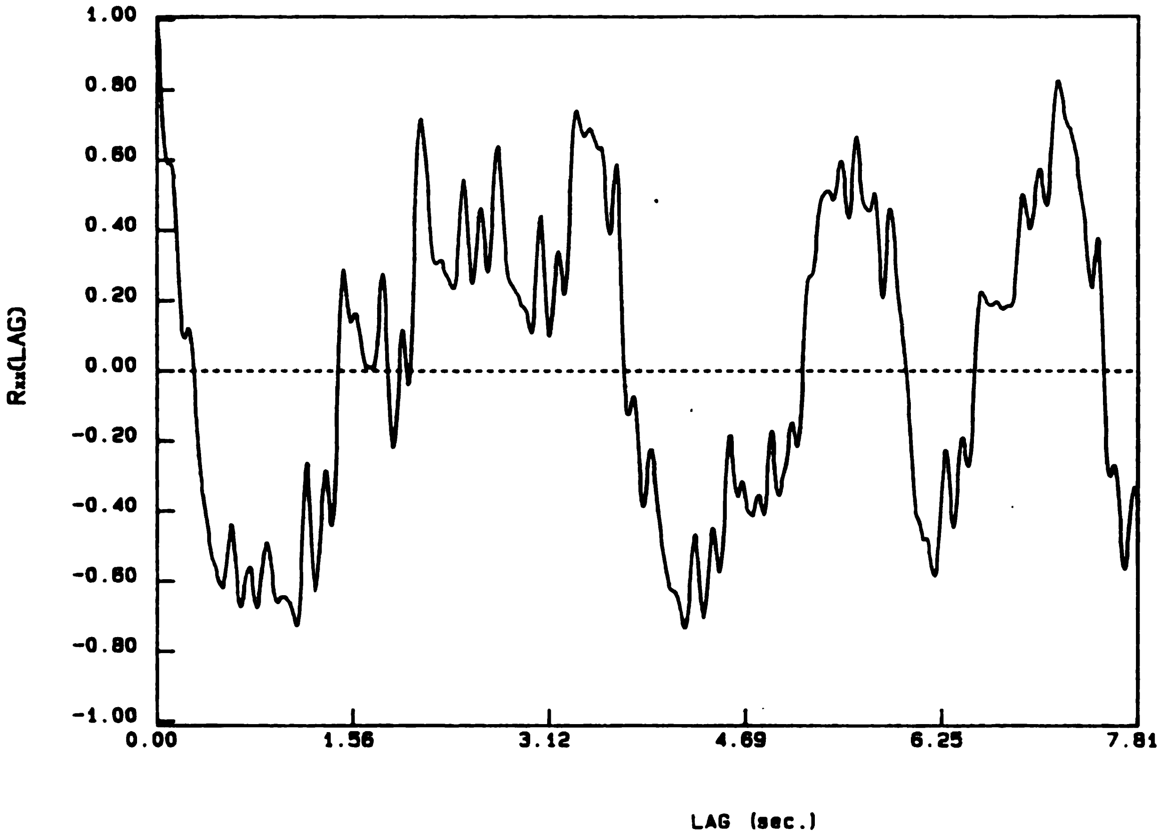

Figure 6b. Autocorrelation of Signal in Figure 6a
(window = 128 points, maximum lag = 1000 points)

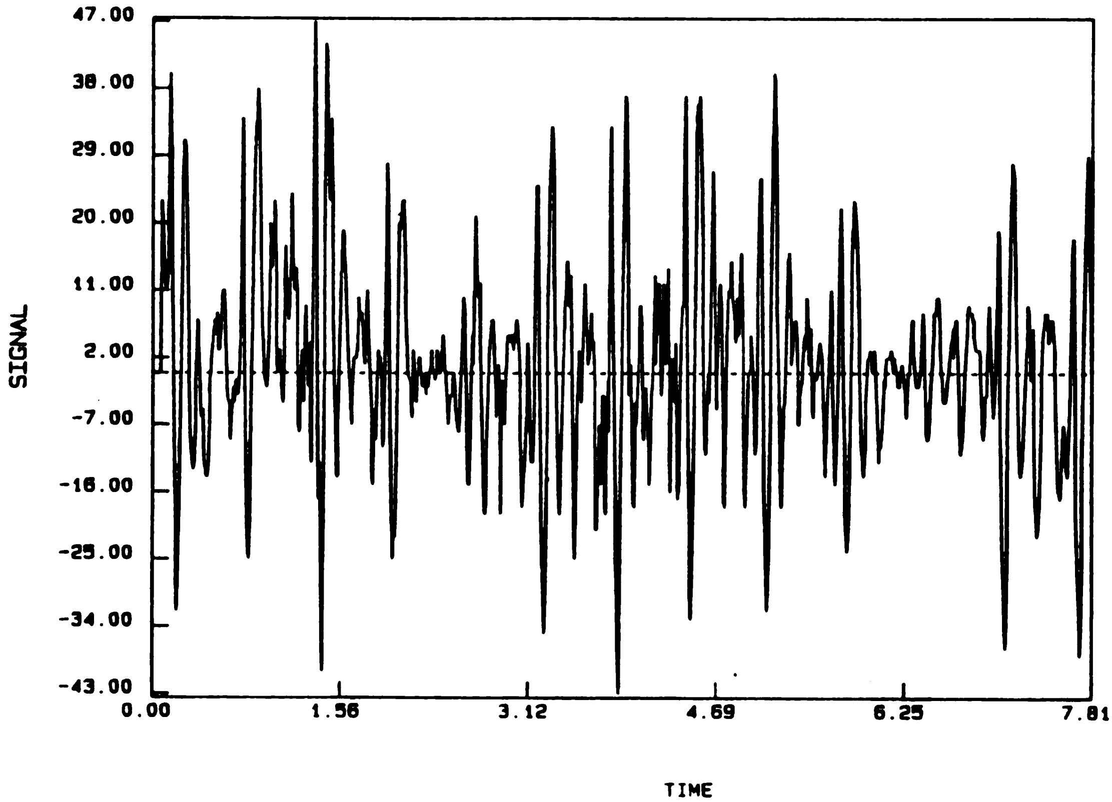

Figure 6c. Prediction Errors of the Simple First Order Predictor, q=.98

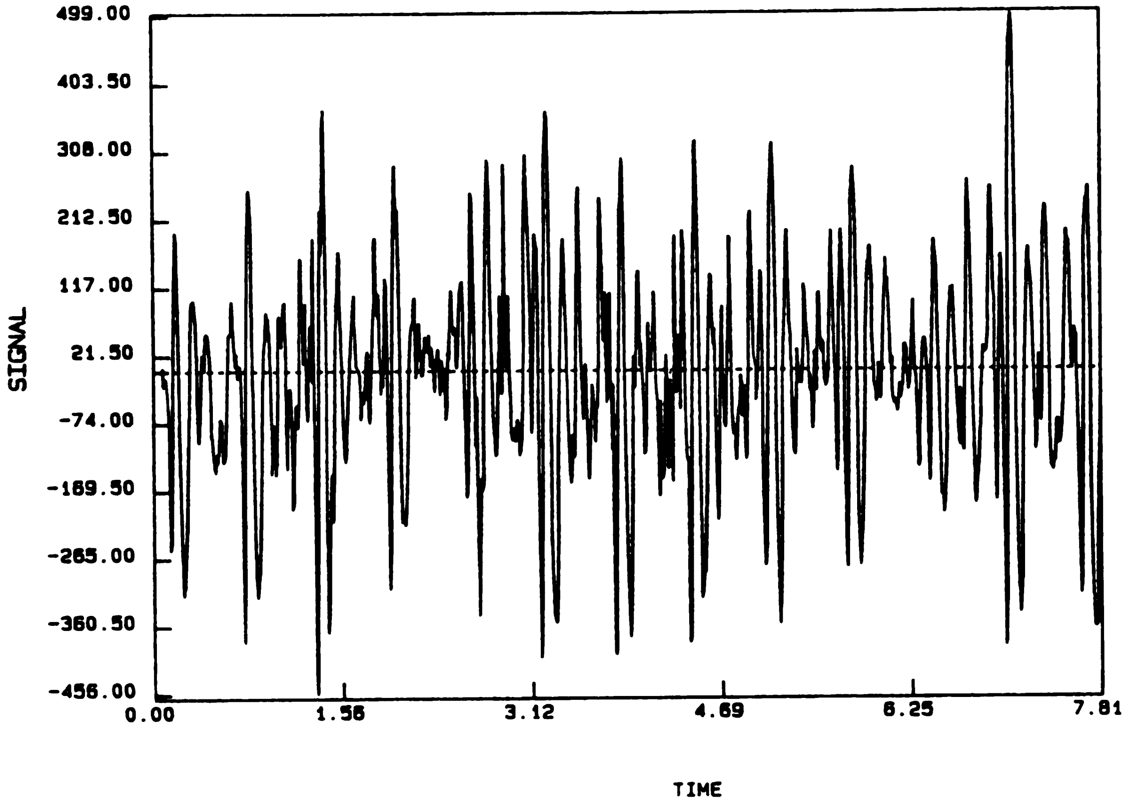

Figure 6d. Prediction Errors of the Normalized Least Squares Lattice filter (p=1,q=.98)

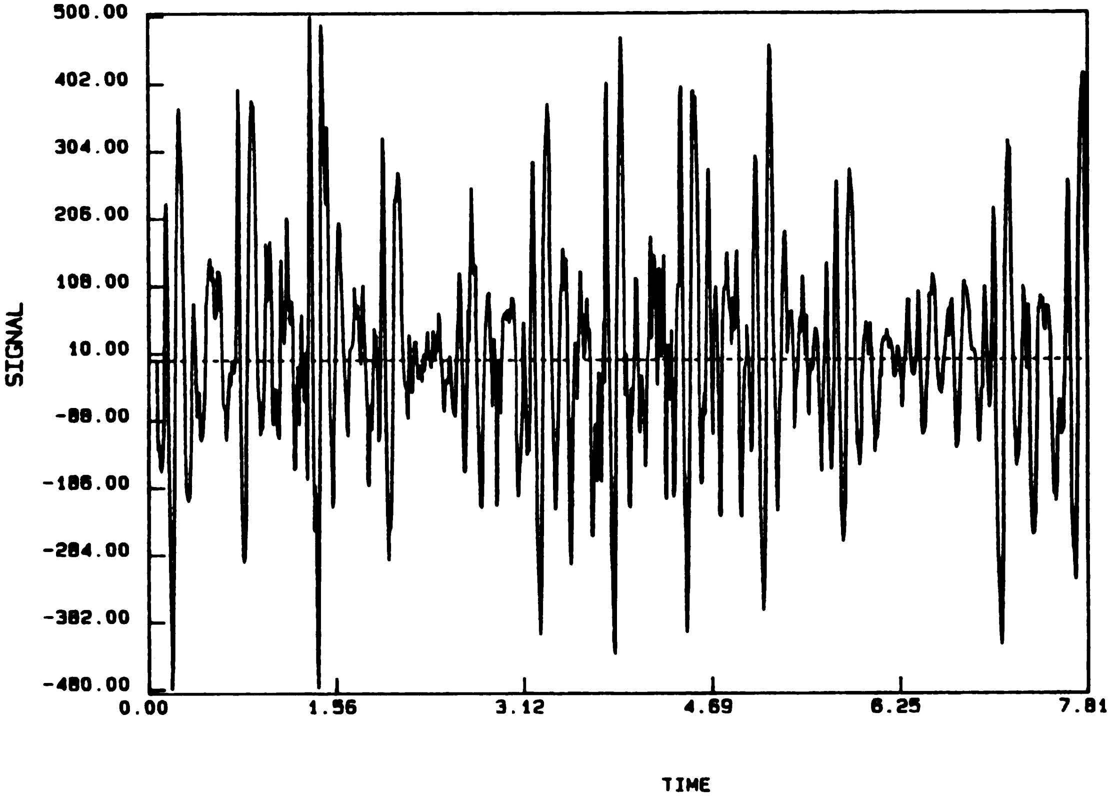

Figure 6e. Prediction Errors of the Pseudo Linear Regression Prediction Algorithm (p=1,q=.98)

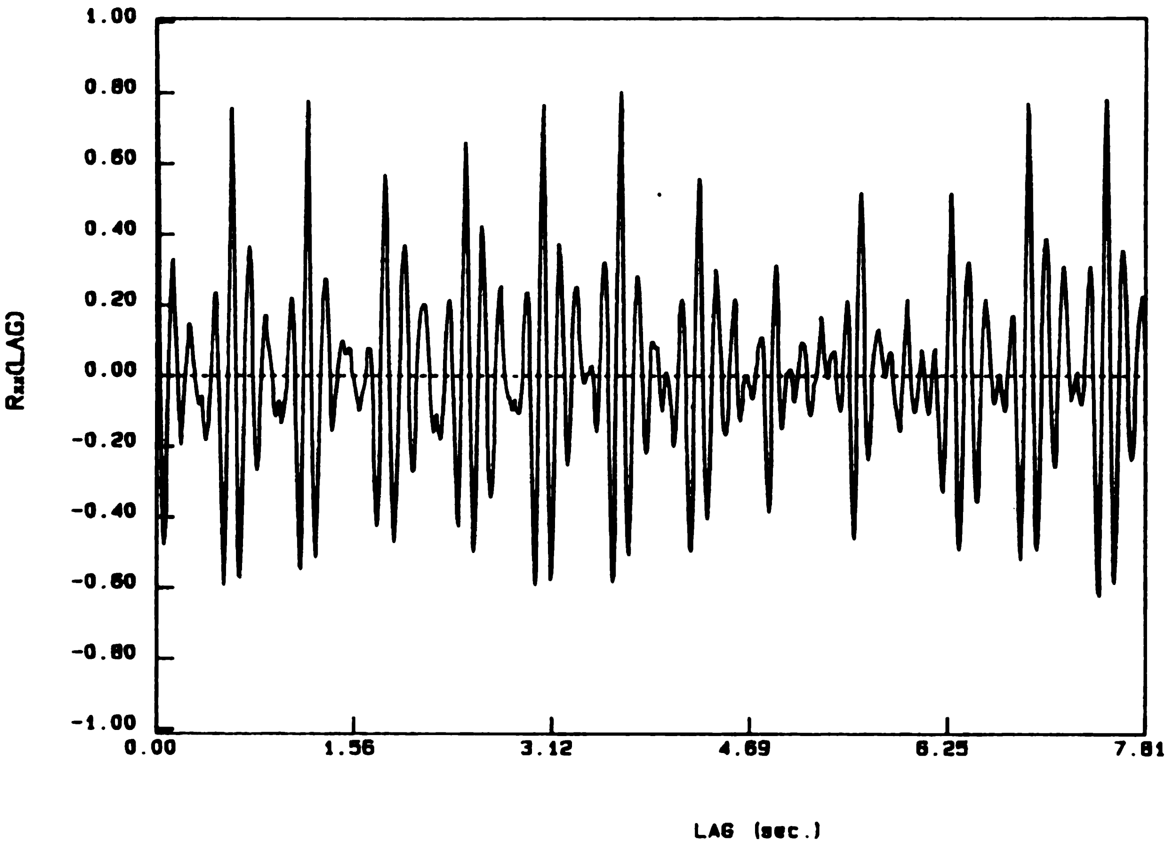

Figure 6f. Autocorrelation of Simple First Order Predictor Prediction Errors (Figure 6c)
(window = 128 points, maximum lag = 1000 points)

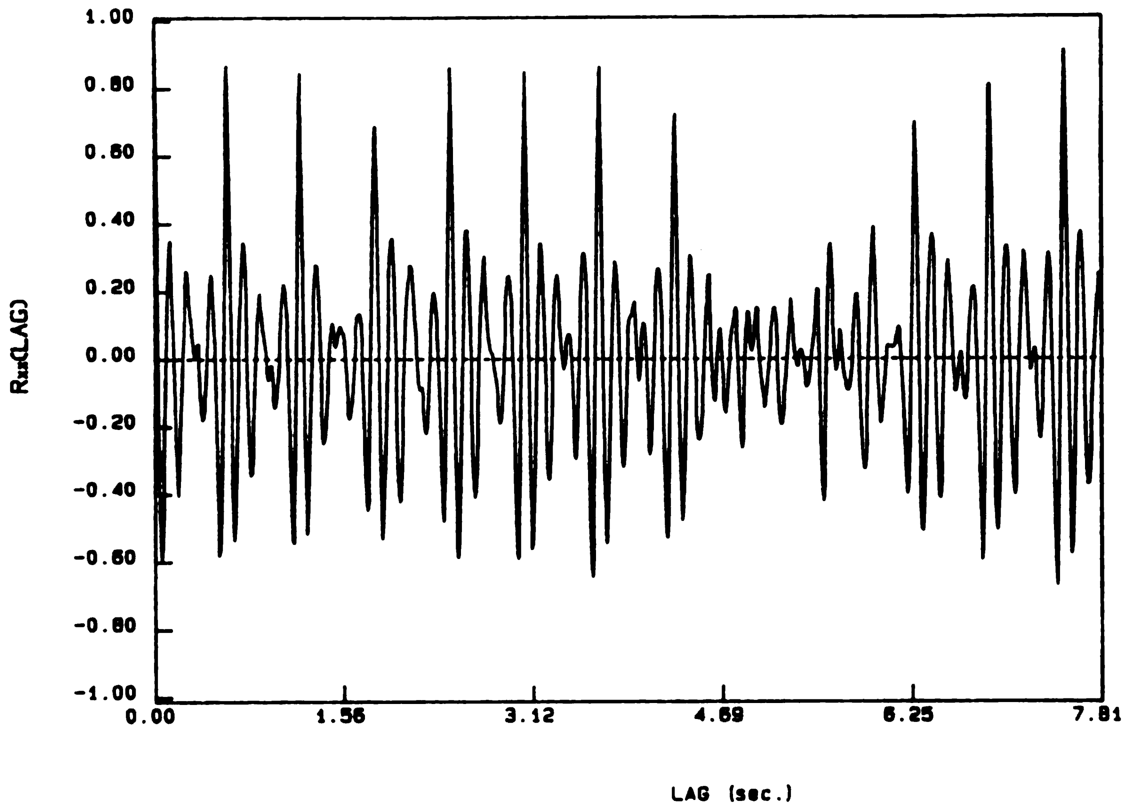

Figure 6g. Autocorrelation of Normalized Least Squares Lattice Filter Prediction Errors (Figure 6d) (window = 128 points, maximum lag = 1000 points)

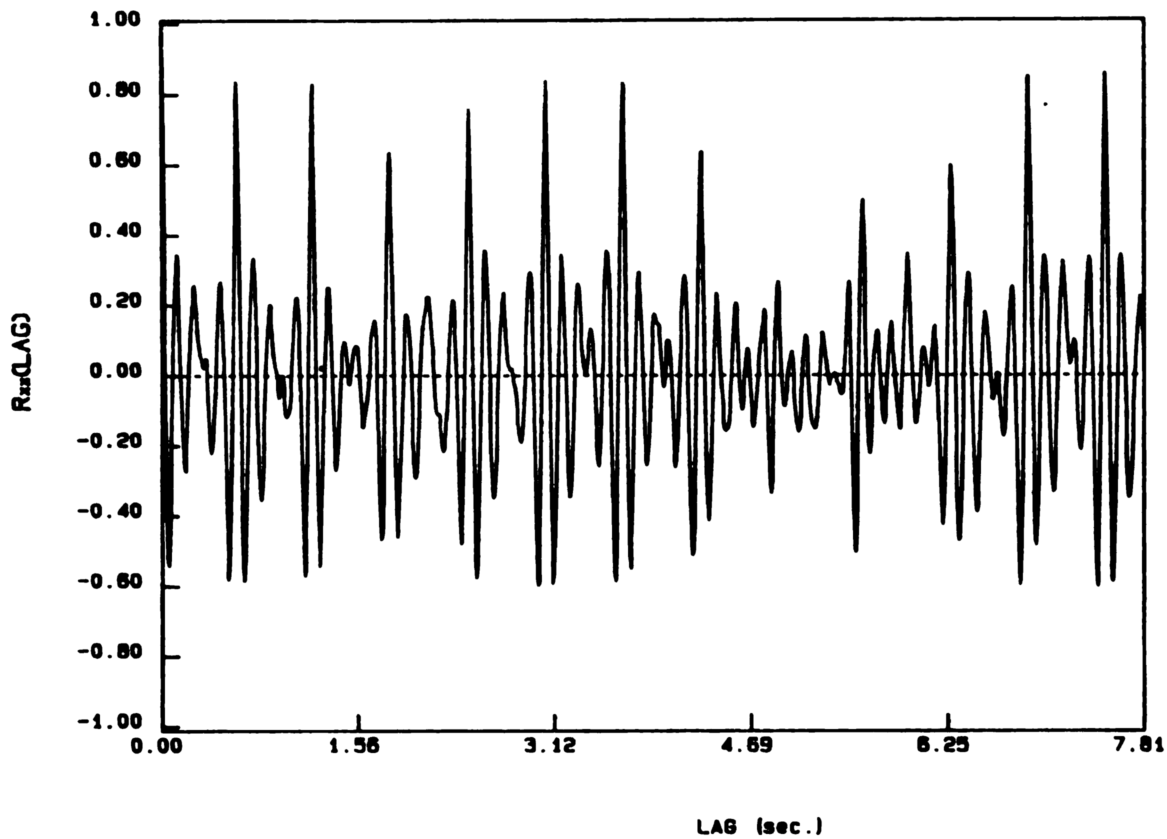

Figure 6h. Autocorrelation of Pseudo Linear Regression Predictor Errors (Figure 6e)
(window = 128 points, maximum lag = 1000 points)

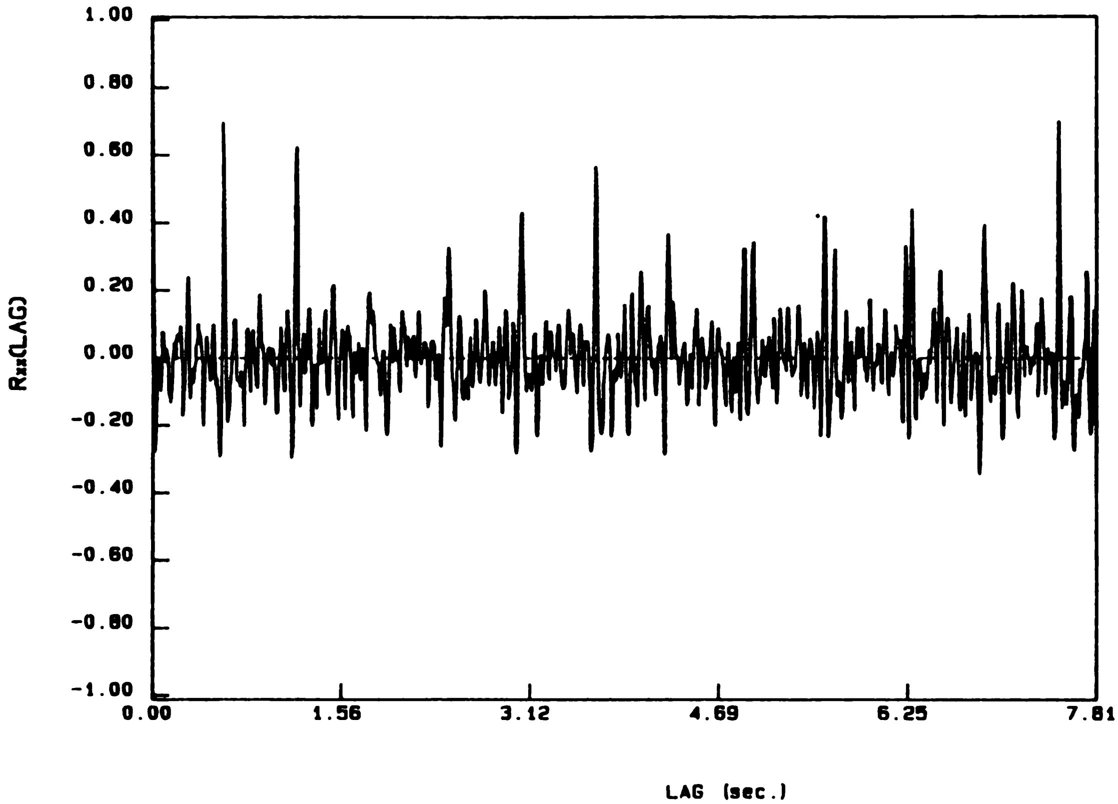

Figure 6i. Autocorrelation of EKG Taken Simultaneously with Microwave Signal
(window = 128 points, maximum lag = 1000 points)

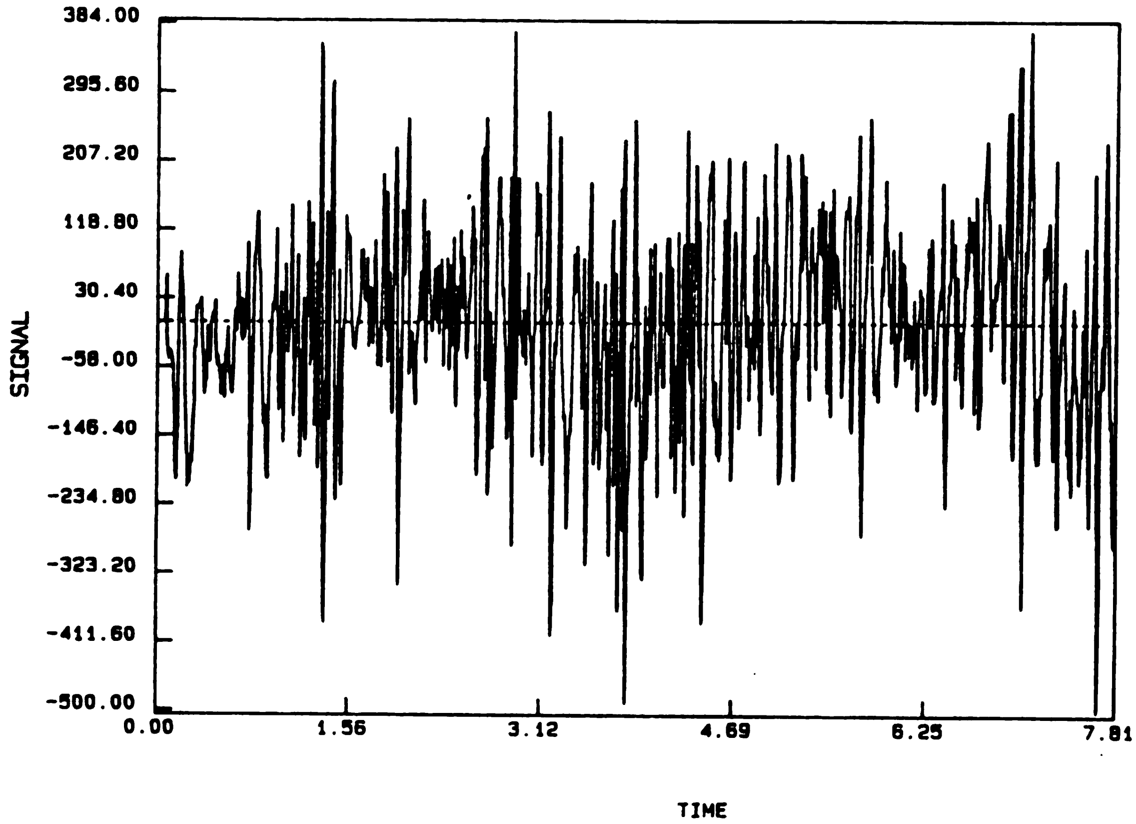

Figure 6j. Prediction Errors of Normalized Least Squares Lattice Filter (p=2,q=.98)

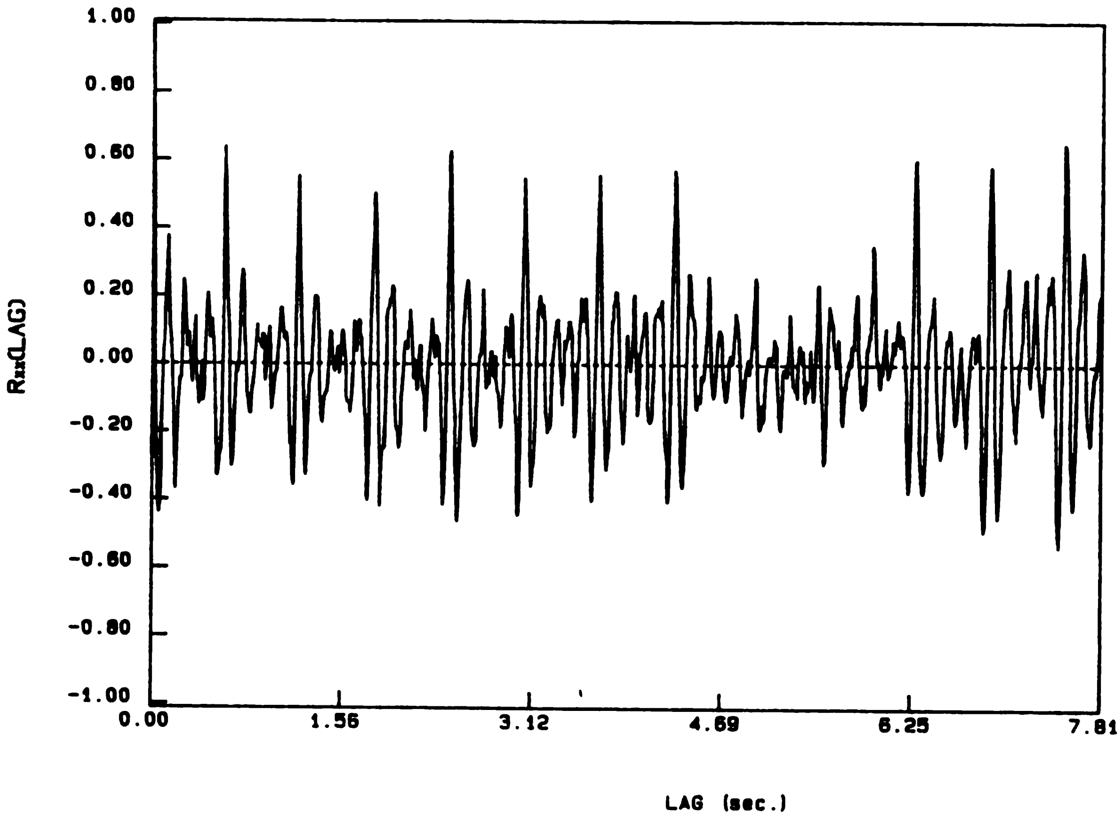

Figure 6k. Autocorrelation of Normalized Least Squares Lattice Filter Prediction Errors (Figure 6j) (window = 128 points, maximum lag = 1000 points)

It is also possible to construct detectors based on statistical parameters computed by the adaptive filters. These have been found to offer the most promise for implementation as heart beat detectors.

The first test criterion is based on an inspection of the prediction errors. If no beat has occurred recently, the model excitation, $e(t) = w(t)$, is normal, zero mean, with variance Rw^2. Given that $e(t)$ is white, the following conditional density describes $e(t)$

$$f(e(t) \mid u(t)=0) = \exp[-(e(t)/Rw)^2 /2)]/\mathrm{sqrt}(2\ Rw^2) \quad (57)$$

A change in $u(t)$ from $u(t-1)$ will be implemented by a log-likelihood ratio test

$$\log[f(e(t) \mid u(t)=0) \ / \ f(e(t) \mid u(t-1)=0)] = -((e(t)/Rw)^2 - (e(t-1)/Rw)^2)/2$$

$$= D1(t) \quad (58)$$

If no heart beat has occured, $D1(t)$ should be small. If $e(t)$ were available, this test could be performed directly. However, under the assumption that the recent model excitation is white, the estimation routines should yield accurate, unbiased estimates of the model parameters. If this is so, the prediction error $e(p,t)$ should be an accurate estimate of the model excitation $e(t)$. Therefore, $(e(t)/Rw)^2$ will be approximated by $(e(p,t)/Re(p))^2$. The variable, $D1(t)$, will be used to mask the prediction error $e(p,t)$, and the occurrence of the heart beat will be chosen as the largest absolute value of $D1(t)e(p,t)$. This test will be used with the simple first order predictor .

A possibility for another detector is to check whether a group of several measurements are jointly Gaussian [4]. This will be the case if the recent model excitation is solely zero-mean Gaussian, i.e. there has been no heart beat. The likelihood function for this event is proportional to

$$f(y(t),...,y(t-k) \mid u(t-i)=0) = \qquad i=1,..k$$
$$\exp[-(y(t),...,y(t-k))W(y(t),...,y(t-k))'/2] \qquad (59)$$

where W is the inverse of the covariance matrix of y. The exponent, g(t), of this likelihood function

$$g(t) = (y(t),...,y(t-k))W(y(t),...,y(t-k))'/2 \qquad (60)$$

is readily available from both the normalized least squares lattice filter (equation (47)) and the pseudo linear regression predictor algorithm (equation (56)).

The above test for jointly Gaussian data is implemented as a log-likelihood ratio test

$$D2(t) = \log[f(y(t),...,y(t-k) \mid u(t-i)=0)/$$
$$f(y(t-1),...,y(t-k-1) \mid u(t-i-1)=0] \qquad (61)$$
$$= g(t) - g(t-1)$$

of the present block of data versus the previous block of data as an indication of a change in u(t) from u(t-1). The block size, k, is equal to the filter order p. The parameter D2(t) is used to mask the residuals e(p,t) in both the normalized lattice and the pseudo linear regression algorithm. The largest value of D2(t)e(p,t) in a specified time interval will be chosen as the location of the heart beat.

The application of the two tests (equations (61) and (58)) will be presented simultaneously. The following steps are performed at each sampling instant. The time series of the intermediate steps are plotted in Figure 7 to show the different behavior of the algorithms. The purpose will be show the performance of the algorithms on relatively uncluttered data.

The first step is to obtain the linear prediction errors, e(p,t), (Figure 7a-c) from the adaptive filters. Then the detector parameters are computed. For the simple predictor, this requires using the autocorrelation estimate (21) in (20) to compute the current estimate of the error variance

$$Re(1,t) = (1-a^2)Ry(0,t) \tag{62}$$

which is then used to compute the term e(1,t)^2/Re(1,t))^2. The present value of this term is then subtracted from the previous value e(1,t-1)^2/Re(1,t-1)^2 to produce D1(t) (Figure 7d). The detector parameters for the normalized lattice and pseudo linear regression algorithm are computed by determining the likelihood function exponent, g(t-1), by equations (48) and (56), and subtracting it from the previous value g(t-2) to produce the parameter D2(t-1) (Figure 7e,f).

Once the detector parameters D1(t) and D2(t-1) are found, they are used to mask the innovations e(p,t) to produce the detector output e(p,t) D(t) (Figure 7d-f). The output is peak detected and the largest value in a specified search range will be chosen as the occurrence of the heart beat. The search range is specified by a maximum and minimum time lag from the previously detected peak. The minimum time lag is the inverse of the minimum expected heart rate and the maximum time lag is the inverse of the maximum heart rate, where both rates are in seconds. This detection can be performed "on the fly" so that the detector outputs D(t) e(p,t) over the entire search range need not be stored.

The time lag between the peak value in the present search range and the peak in the previous search range is used to generate the estimate of the instantaneous heart rate.

Table 1 shows side by side the results of the four approaches: peak detection, the simple algorithm, the normalized lattice algorithm, and the pseudo linear regression algorithm. The inclusion of EKG data allows for quantitative comparison. The normalized lattice detector is clearly superior to the other methods. For this signal, the variance in the BPM estimates is decreased from about 6 BPM to about 1 BPM.

The previous results show a marked improvement in BPM estimate using the normalized lattice filter. The following results will show that this form of detection can yield results in situation where peak detection fails completely and the only competetive form of processing is autocorrelation with large window sizes.

The data presented in Figure 8a was taken at a distance of three feet from a seated subject who had just performed fifty push-ups. This combination of measurement from a distance and chest movement due to heavy breathing obscures the heart signal. Inspection of the signal shows that peak detection is inapplicable in this case. Figure 8 presents the detector outputs from the three filtering techniques. Of these techniques, the normalized lattice is the only one which produces reliable results (Table 2). Only one beat was not detected. There are several possible reasons for this superior behavior. When compared to the simple predictor, the normalized lattice is better able to track non-stationary data due to the likelihood variable which improves tracking behavior. That the data is highly non-stationary is further suggested in that the optimum detector performance was obtained for q = .91, a fairly low value for the forgetting factor, which means that the data is stationary over only short intervals. The pseudo linear regression method also contains this likelihood variable, however, the parameters in the normalized lattice algorithm are normalized by their standard deviations. This form of normalization, which is a test for non-zero mean Gaussian outliers [28], accentuates the occurence of the heart beat.

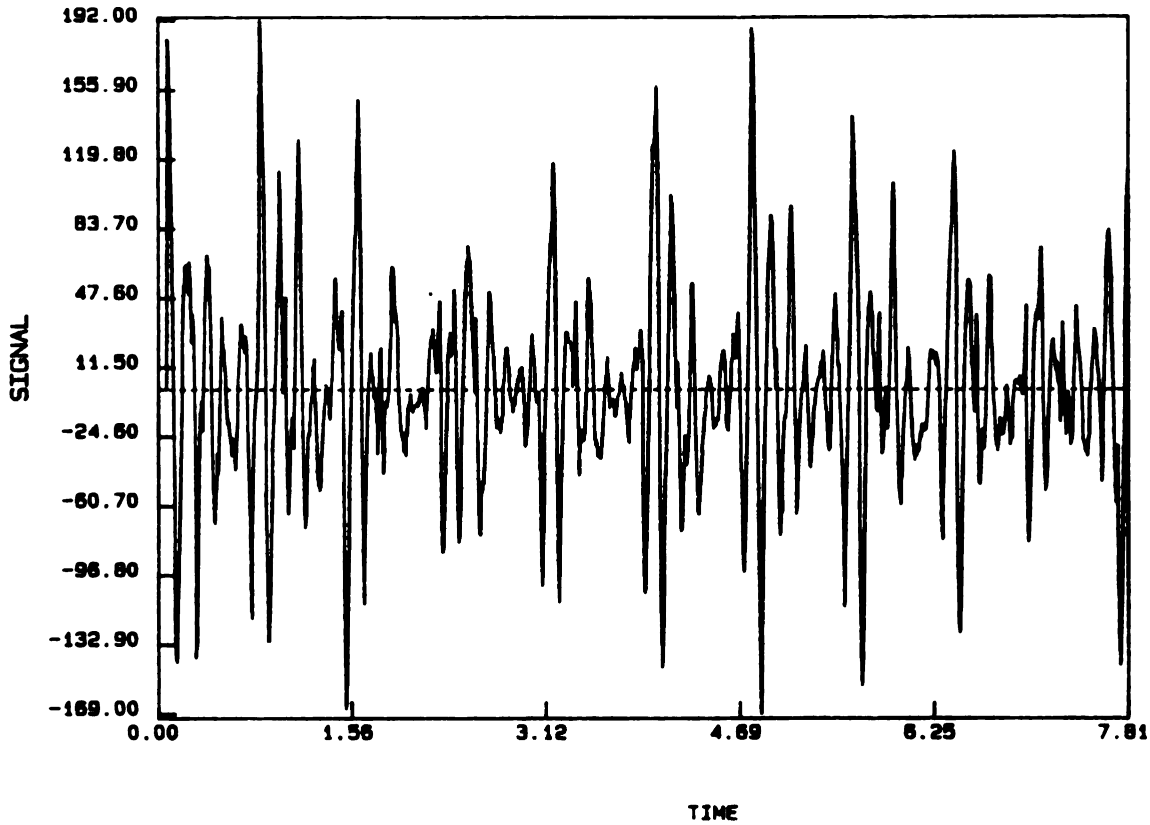

Figure 7a. DI(t) of Simple First Order Predictor for Data in Figure 5a., q = .98

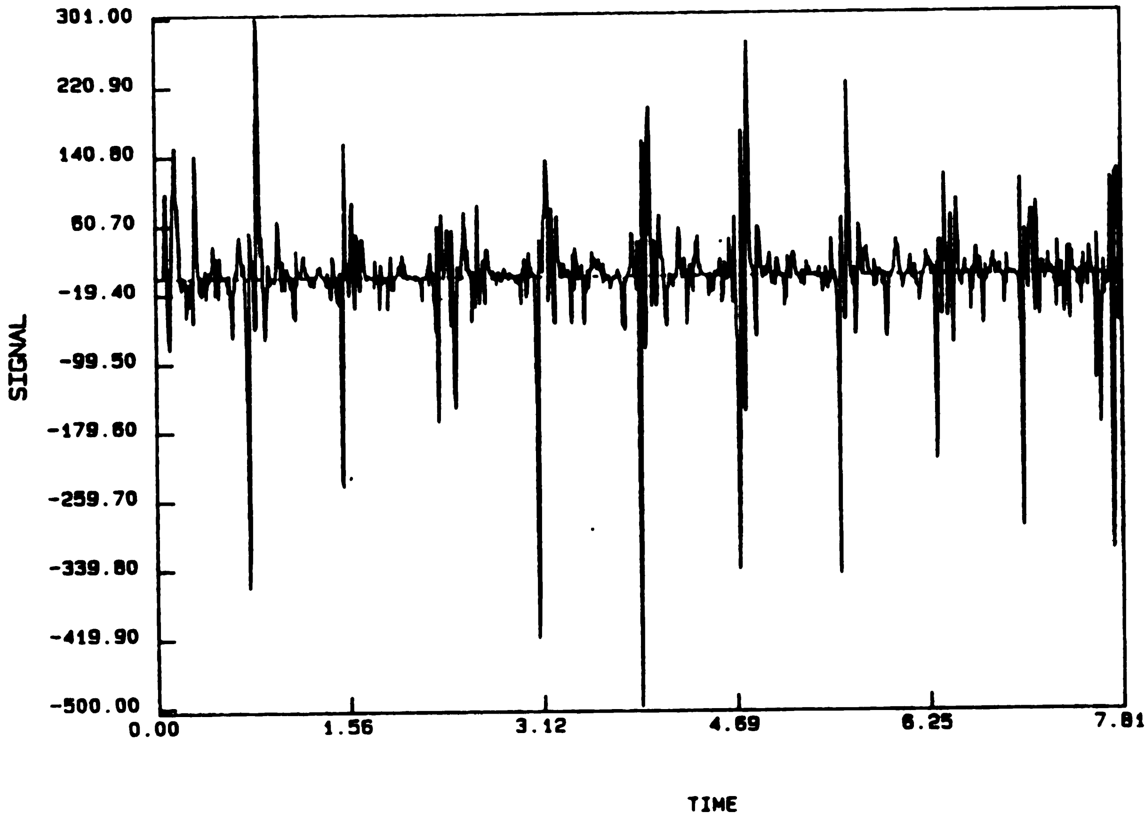

Figure 7b. D2(t) of Normalized Least Square Lattice Algorithm of Data in Figure 5a, p=1, q=.98

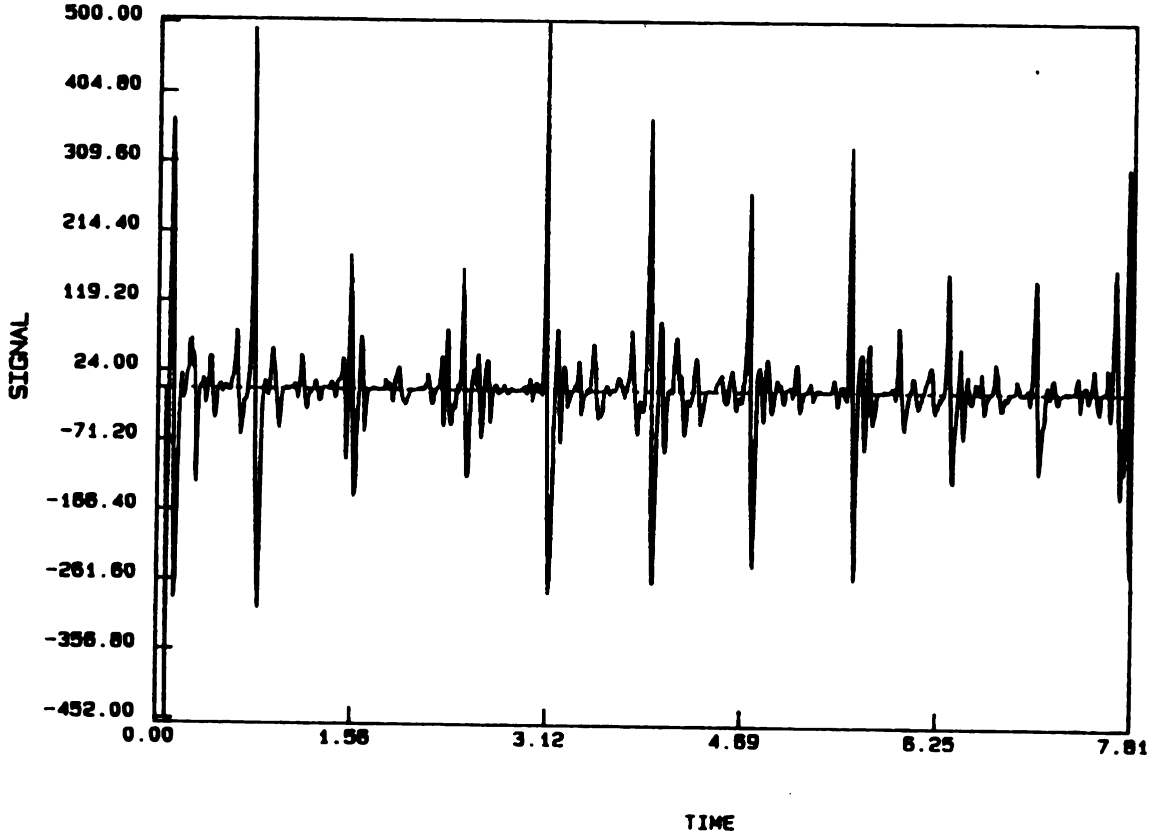

Figure 7c. D2(t) of Pseudo Linear Regression Predictor for Data in Figure 5a, p=1,q=.98

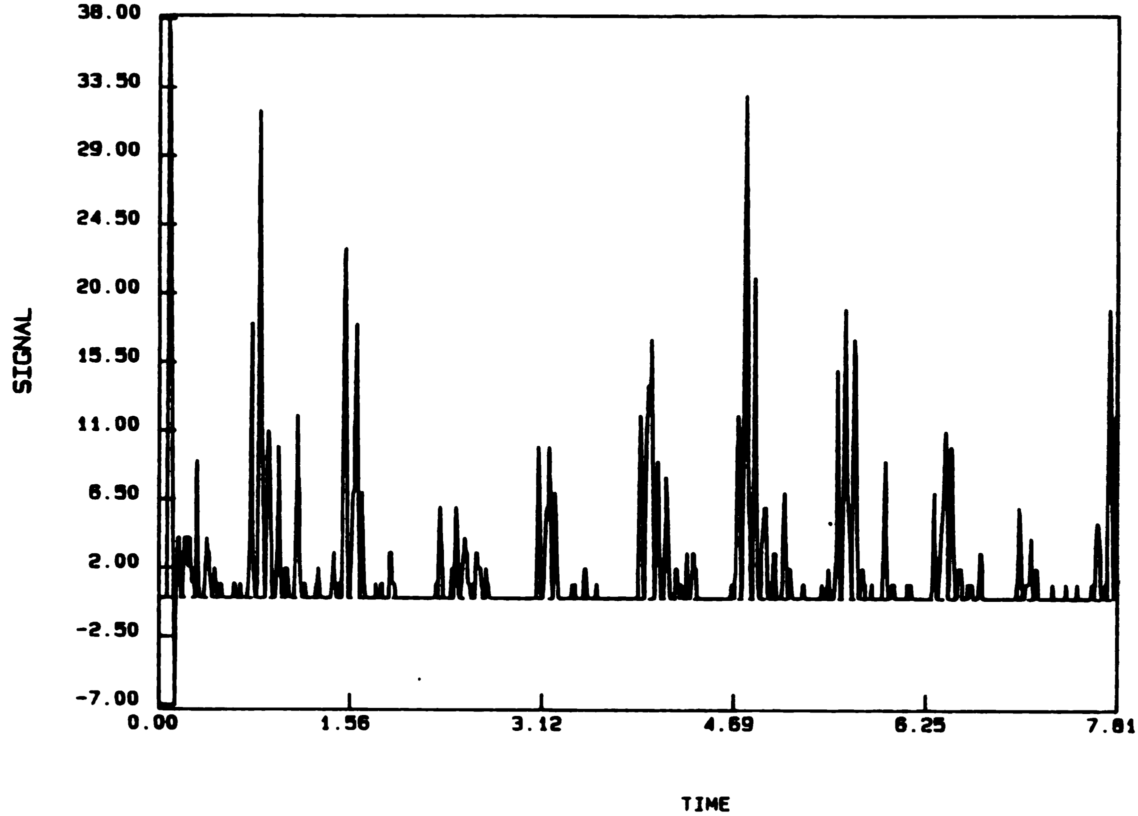

Figure 7d. Detector Output, D1(t) e(1,t), of Simple First Order Predictor for Data in Figure 5a, q=.98

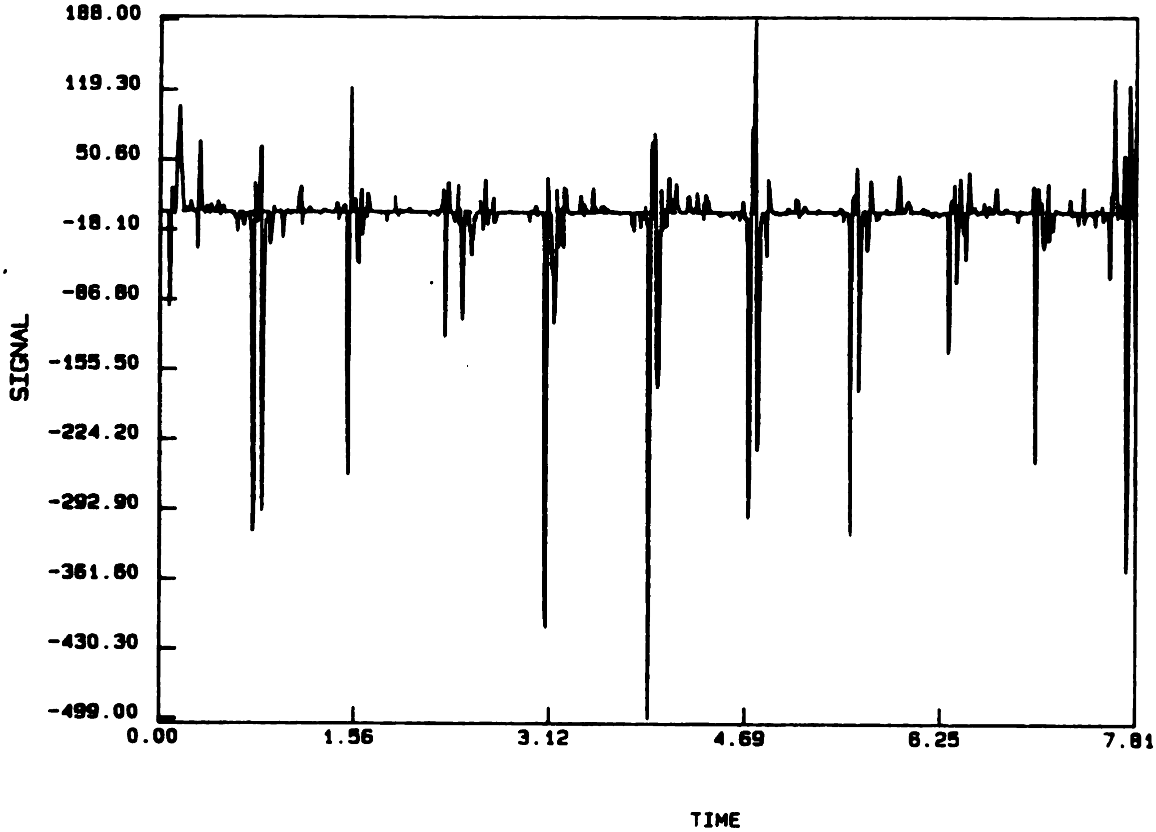

Figure 7e. Detector Output, D2(t) e(p,t), of Normalized Least Squares Lattice Filter for Data in Figure 5a, p=1,q=.98

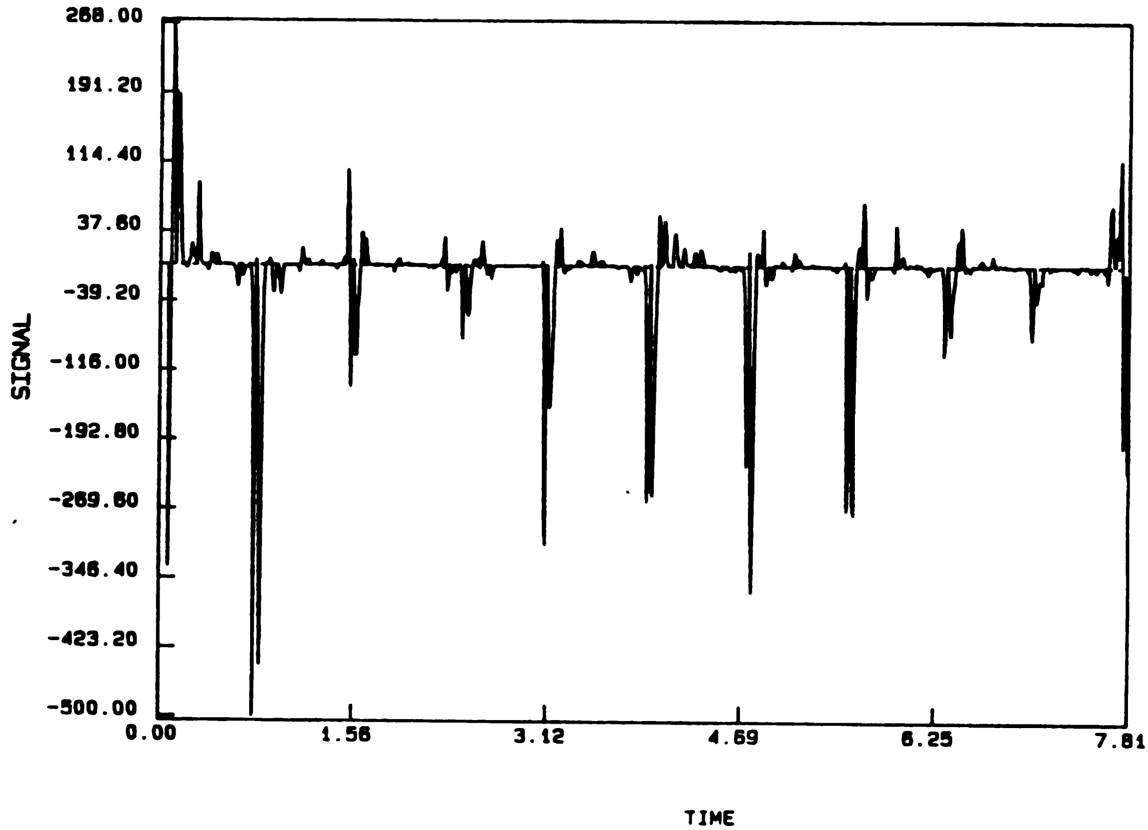

Figure 7f. Detector Output, D2(t) e(p,t), of Pseudo Linear Regression Predictor for Data in Figure 5a, p=1,q=.98

Table 1. Detector Performance on Data in Figure 7
Instantaneous Heart Beat Measurements (BPM)

Filter Order = 1, q = 0.98

EKG (reference)	Peak Detection	Simple First Order Predictor	Normalized Least Squares Lattice	Pseudo Linear Regression Predictor
79.17525	78.36735			
77.57571	66.78266	66.20689	77.57570	67.36842
74.56315	89.30217	90.35287	74.56306	89.30232
73.14281	73.14270	71.77574	73.14290	69.18922
75.29412	75.29412	71.11111	75.29412	75.29412
73.14290	73.84615	73.84615	73.84615	73.84615
77.57571	76.80000	82.58059	76.80000	80.84206
85.33334	85.33334	86.29219	85.33334	86.29219
82.58070	81.70213	80.84206	82.58059	77.57571
78.36735	79.17525	73.84620	79.17531	78.36739
74.56311	74.56311	74.56311	73.84615	74.56311
62.43903	62.43890	57.31317	61.93548	65.08441
57.74452	55.25165	66.78317	58.18182	58.18241
70.45808	74.56381	66.20689	69.81779	66.78239
74.56381	73.84615	73.84615	75.29457	74.56284
76.03988	76.03887	68.57143	75.29367	79.17556
85.33264	86.29249	98.46077	86.29249	81.70160
87.27273	87.27273	87.27273	87.27273	87.27273
85.33276	84.39526	85.33392	85.33392	85.33392
		87.27285	80.84180	

Error Variance: 32.76 139.6 0.3287 22.28
(BPM**2)

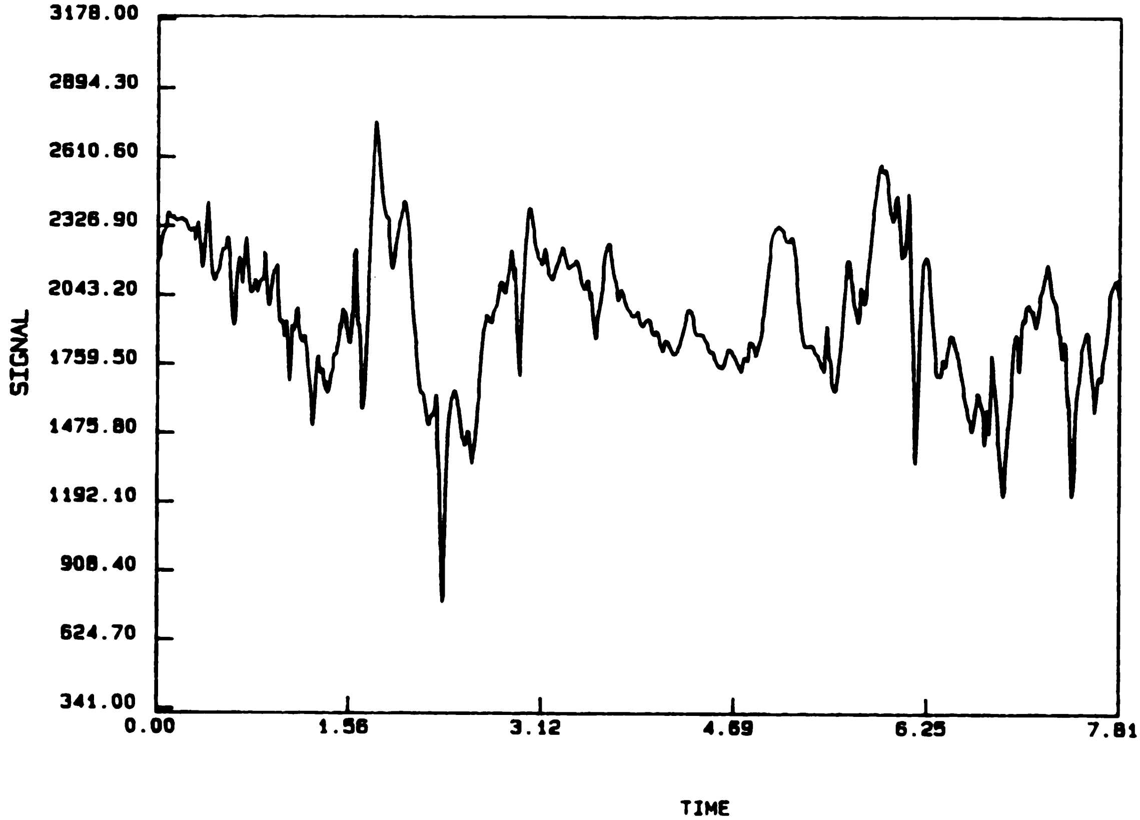

Figure 8a. Microwave Returns from an Exercised Subject Seated Three Feet from the Microwave Transceiver (sampling rate = 128 Hz)

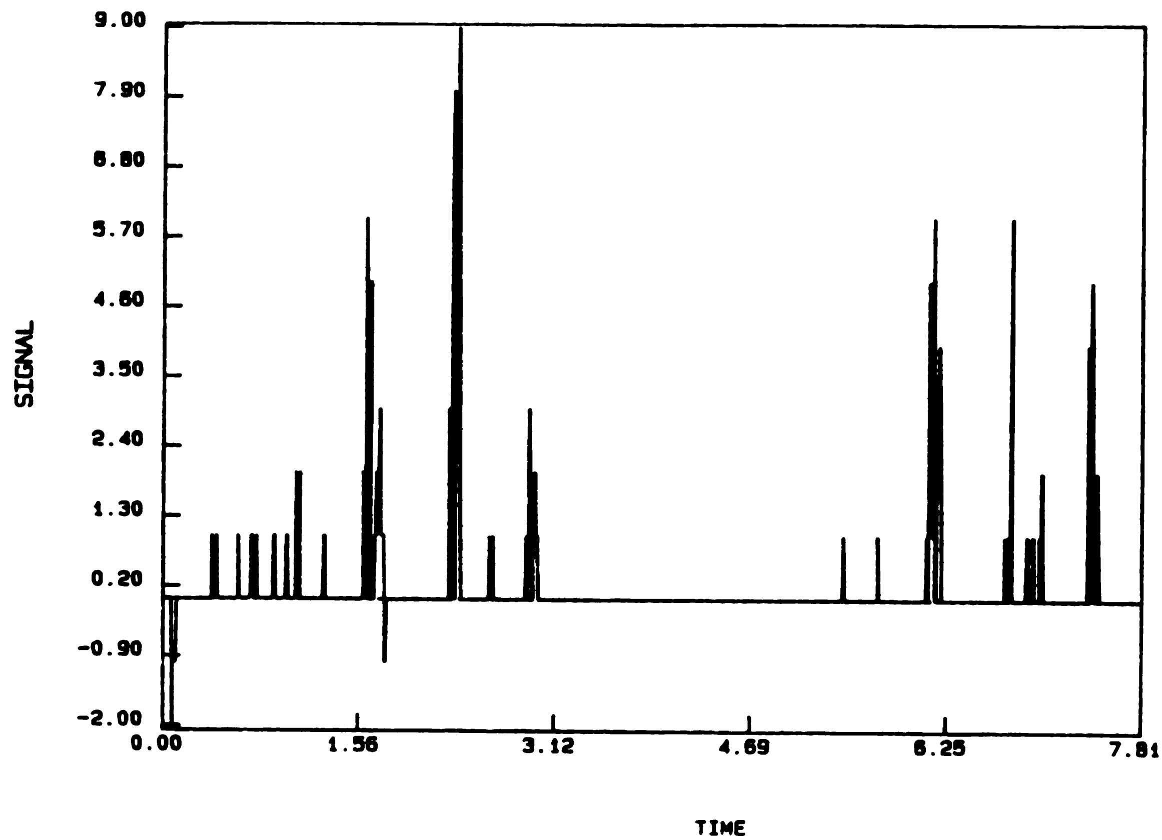

Figure 8b. Detector Output, DI(t) e(1,t), of Simple First Order Predictor, q=.91

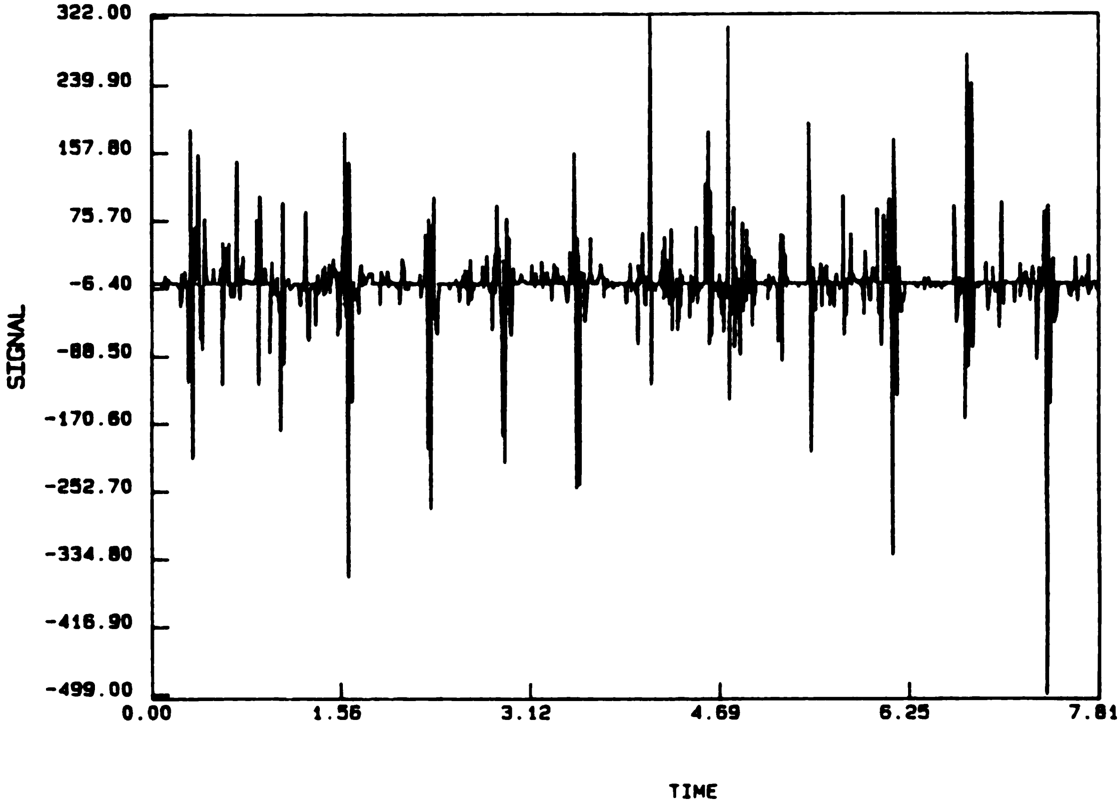

Figure 8c. Detector Output, D2(t) e(p,t), of Normalized Least Squares Lattice Filter, p=1,q=.91

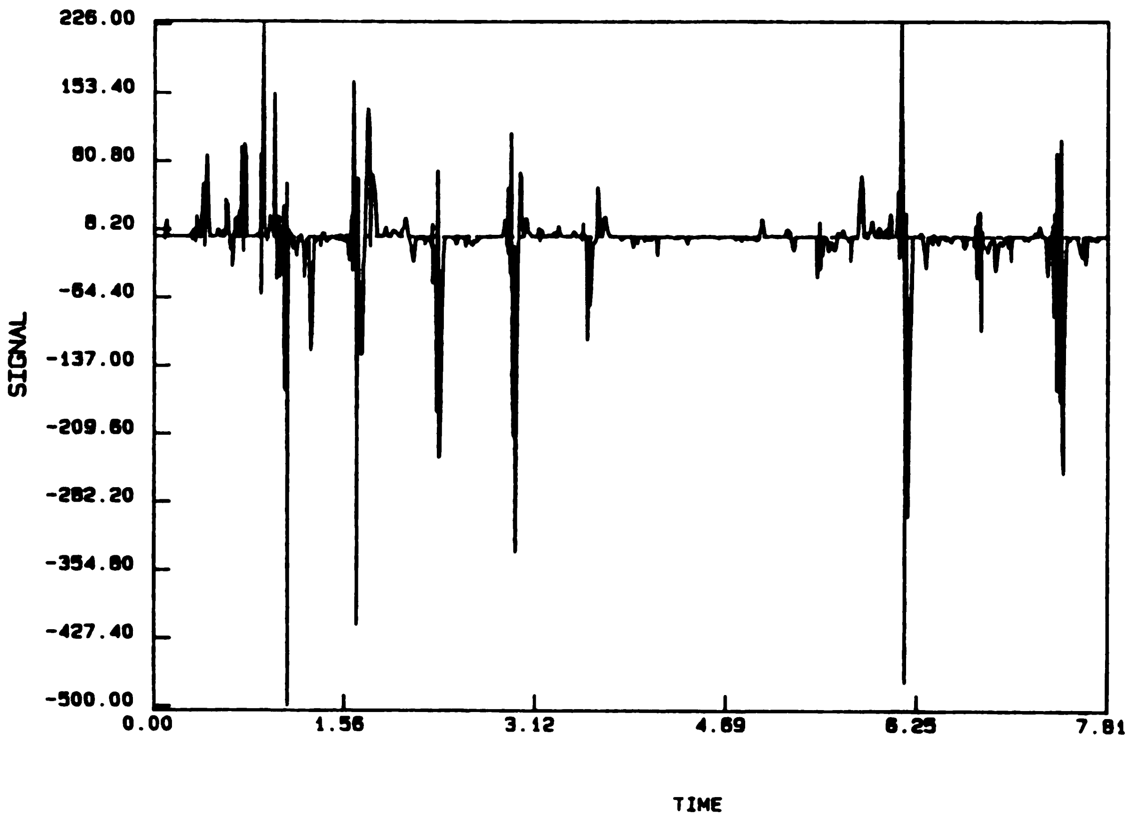

Figure 8d. Detector Output, D2(t) e(p,t), of Pseudo Linear Regression Predictor p=1,q=.91

Table 2. Detector Performance on Data in Figure 8

Filter Order = 1, q = .91

Instantaneous Heart Beat Measurements (BPM)

	EKG (reference)	Normalized Least-Squares Lattice
	92.53012	
	89.30232	88.27592
	99.74018	98.46154
	97.21526	102.3999
	94.81474	98.46154
	92.53019	92.53019
	90.35307	88.27580
	93.65833	89.30232
	94.81489	97.21526
	94.81474	92.53026
	80.84216	83.47815
	58.18182	59.07692
	70.45879	67.96460
	82.58054	82.58097
	88.27561	90.35191
	93.65923	96.00000
	91.42857	64.00000
Error Variance (BPM**2)		33.04

CONCLUSION

It has been shown that application of adaptive filtering to remote heart rate monitors improves heart rate measurements. The filters can either be employed as pre-processing for autocorrelation or peak detection or they can be used alone as detectors. In particular, evidence has been presented which indicates that a detector based on the normalized, least squares lattice filter is the superior choice among detectors investigated to date; it is faster, less complex and more accurate than any other algorithms investigated.

POSSIBILITIES FOR FURTHER INVESTIGATION

The application of the adaptive FIR filters to this problem was justified by the signal model given in equation (1). It is possible that other models are more appropriate, which would permit the use of other forms of adaptive filters. One possibility is to use a pole-zero, or autoregressive moving average (ARMA) model for the signal. The appropriate adaptive filter would then have both poles and zeros [29]. The key to all approaches is to find a model for the signal. A possible approach might be to use a frequency domain approach, with model poles and zeros chosen to match the signal spectrum [30]. Parametric identification routines could then be used to form an adaptive filter, as in [31]. This approach has several potential benefits. It is easy to constrain the model poles to remain within the unit circle (although this does not guarantee stability in the time-varying case). It also allows modeling the signal spectrum as narrowband components, which is a good description of the heart signal spectrum. Also, with a general model form, many of the identification techniques presented in [22] can be applied.

Clutter poses major problems in all detection schemes. Models for the different types of clutter, such as mechanical vibration and moving trees, could be developed. Adaptive filters based on a combined model of heart signal and clutter could then be implemented.

Finally, it might be desirable to find a way to introduce past heart rate measurements into the filtering process. This would require modeling how the heart signal varies with heart rate.

APPENDIX

APPENDIX

NORMALIZED LATTICE FILTER RECURSIONS

The projection operator is defined

$$P\{Yp(t)\} = Yp(t)(Yp(t)'\ Yp(t))^{-1}\ Yp(t)'$$

Inserting the the least squares solution (18) into the prediction error equation (15)

$$\begin{bmatrix} e(p,0) \\ e(p,1) \\ \vdots \\ \vdots \\ \vdots \\ e(p,t) \end{bmatrix} =$$

$$\begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ \vdots \\ \vdots \\ y(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 & \dots & 0 \\ y(0) & 0 & \dots & \vdots \\ y(1) & y(0) & \dots & 0 \\ \vdots & & & y(0) \\ \vdots & & & \vdots \\ y(t-1 & & & \dots y(t-p) \end{bmatrix} (Yp(t)'\ Yp(t))^{-1}\ Yp(t)'\ y(0:t)$$

$$= (I - Yp(t)(Yp(t)'\ Yp(t))^{-1}\ Yp(t)'\ )\ y(0:t) \qquad (A1)$$

The error e(p,t) can be extracted using a matrix operation on (A1). Substituting the projection operator definition into (A1)

$$e(p,t) = [0 \dots 0\ 1]\ (I - P\{Yp(t)\})y(0:t) \qquad (A2)$$

$$= h'\ (I - P\{Yp(t)\})\ y(0:t)$$

$$= h' (I - P\{Yp(t)\})) (I - P\{Yp(t)\}) y(0:t)$$

where $h' = [0...0\ 1]$.

To normalize e(p,t) by its standard deviation, an analogy will be drawn with the covariance of two zero mean random variables

$$r(x,y) = cov(x,y) = E(xy)/sqrt(E(x^2)E(y^2)) \qquad (A3)$$

For an expression in the form of equation (A4), $V'[I-P\{s\}]W$, the normalization will be

$$r\{s\}(V,W) = [V'(I-P\{s\})(V'(I-P\{s\})']^{-1/2} V'[I-P\{s\}]W * [((I-P\{s\})W)'(I-P\{s\})W]^{-1/2}$$

$$= [V'(I-P\{s\})V]^{-1/2} V'[I-P\{s\}]W [W'(I-P\{s\})W]^{-1/2} \qquad (A4)$$

using the fact that $(I-P\{y(p,t)\})$ is symmetric and idempotent.

Using this notation, the normalized prediction error, e~(p,t), can be written

$$e\tilde{}(p,t) = r\{s\}(h,y(0:t)) , s = Yp(t) \qquad (A5)$$

Similarly, the backward prediction errors can be put into normalized form. The coefficients of the backward error filter are determined in the same way as the coefficients of the forward filter. Both predictions involve a projection onto Yp(t). The derivation of the backward predictor proceeds from the definition of the backward prediction error

$$r(p,t-1) = y(t-p-1) + b(p,p)\, y(t-1) + \ldots + b(1,p)\, y(t-p) \quad \text{(A6)}$$

$$\begin{bmatrix} 0 \\ r(p,0) \\ : \\ : \\ : \\ : \\ : \\ r(p,t-1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ : \\ : \\ 0 \\ y(0) \\ : \\ y(t-p-1) \end{bmatrix} + \begin{bmatrix} 0 & \ldots & 0 \\ y(0) & & : \\ y(1) & & : \\ : & & : \\ y(p-1) & & y(0) \\ y(p) & & y(1) \\ : & & : \\ y(t-1) & \ldots & y(t-p) \end{bmatrix} \begin{bmatrix} b(p,p) \\ : \\ : \\ : \\ : \\ : \\ : \\ b(1,p) \end{bmatrix} \quad \text{(A7)}$$

The least squares solution for the backward predictor coefficients is obtained as

$$\begin{bmatrix} b(p,p) \\ : \\ : \\ : \\ : \\ b(1,p) \end{bmatrix} = -\,(Yp(t)'\ \ Yp(t))\hat{\ }{-1}\ Yp(t)\ ' \begin{bmatrix} 0 \\ : \\ 0 \\ y(0) \\ : \\ y(t-p-1) \end{bmatrix} \quad \text{(A8)} \quad \text{(p zeros)}$$

By inserting the least squares solution into the prediction error equation and pre-multiplying by [0..0 1], the backward prediction error can be extracted

$$r(p,t-1) =$$

$$[0\ldots0\ 1][I - Yp(t)(Yp(t)'\ Yp(t))\hat{\ }{-1}\ Yp(t)'\ ]\ y(0{:}t)p{+}1$$

$$= [0\ldots0\ 1][I - P\{Yp(t)\}]\ y(0{:}t)p{+}1 \quad \text{(A9)}$$

where $y(0:t)_{p+1} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ y(0) \\ \vdots \\ y(t-p-1) \end{bmatrix}$

To express the backward prediction errors $r(p,t)$ in their normalized form, $r^{\sim}(p,t)$, the notation will be

$$r^{\sim}(p,t) = r\{s\}(h, y(0:t)_{p+1}) \; , \; s = Y_p(t) \qquad \text{(A10)}$$

In this derivation, the reflection coefficients will be assumed equal. In other forms, such as the unnormalized pre-windowed form [16], $k_r(p+1)$ is computed seperately from $k_e(p+1)$. In the time average sense, the two should be equal, but their instantaneous values are not necessarily equal. By assuming equality, however, the number of operations is reduced.

The reflection coefficient was defined in equation (29) as $k_e(p+1) = -L_r(p+1)/R_e(p)$ and $k_r(p+1) = L_e(p+1)/R_r(p)$. Dropping the notation of r and e on K,

$$k(p+1,t) = L(p+1,t)/\mathrm{sqrt}(R_e(p)R_r(p)) \qquad \text{(A11)}$$

where $L(p+1,t) = E\{e(p,t)r(p,t-1)\}$, the cross correlation of the forward and backward prediction errors at lag 1. $L(p+1,t)$ will be replaced by a time average approximation

$$L(p+1,t) = [0 \; r(p,0) \; \ldots \; r(p,t-1)] \begin{bmatrix} e(p,0) \\ \vdots \\ e(p,t) \end{bmatrix} \quad \text{(A12)}$$

$$= [(I-P\{y(p,t)\})y(0:t)p+1]' \; [(I-P\{y(p,t)\}) \; y(0:t)]$$

$$= y(0:t)p+1' \; (I-P\{y(p,t)\}) \; y(0:t)$$

The error variance Rr(p) can be also be approximated by a time average

$$Rr(p) = [0 \; r(p,0) \; \ldots \; r(p,t-1)] \; [0 \; r(p,0) \; \ldots \; r(p,t-1)]'$$

$$= [\; (I-P\{y(p,t)\}) \; y(0:t)p+1 \;]' \; [\; (I-P\{y(p,t)\}) \; y(0:t)p+1 \;]$$

$$= y(0:t)p+1' \; (I-P\{y(p,t)\}) \; y(0:t)p+1 \quad \text{(A13)}$$

Similarly, Re(p) can be expressed

$$Re(p) = y(0:t)' \; ((I-P\{y(p,t)\}) \; y(0:t) \quad \text{(A14)}$$

It appears that the reflection coefficient can be expressed as the covariance between e(p,t) and r(p,t) at lag 1. Using the equations (A11)-(A14), the approximation of the reflection (A11) can be expressed in the form of equation (A4)

$$k(p+1,t) = [y(0:t)' \; ((I-P\{y(p,t)\}) \; y(0:t)]^{-1/2} \; *$$

$$[y(0:t)p+1' \; (I-P\{y(p,t)\}) \; y(0:t)] \; *$$

$$[y(0:t)p+1' \; (I-P\{y(p,t)\}) \; y(0:t)p+1]^{-1/2}$$

$$= r\{s\}(y(0:t)p+1,y(0:t)), \quad s = Yp(t) \quad \text{(A15)}$$

$$= r\{s\}(y(0:t),y(0:t)p+1)), \quad \text{due to symmetry}$$

Now, recursions will be developed for higher order predictors in terms of lower order predictors and present predictors in terms of past predictors. The first recursion finds $r\{s+x\}$, when $\{s+x\}$ indicates a change in the projection space from $\{s\}$. This will correspond to either an order update, in which $Yp(t)$ becomes $Yp{+}1(t)$, or a time update, when $Yp(t)$ becomes $Yp(t+1)$.

The first step is to note that the space spanned by $\{s+x\}$ is the same as that spanned by $\{s + (I-P\{s\})x\}$, where $P\{s\}$ is the projection operator onto the space $\{s\}$. Therefore,

$$P\{s+x\} = P\{s\} + P\{(I-P\{s\})x\}$$

and

$$I - P\{s+x\} = I - P\{s\} - P\{(I-P\{s\})x\} \qquad \text{(A16)}$$

Using the least squares solution for the projection,

$$P\{(I-P\{s\})x\} = (I-P\{s\})x([(I-P\{s\})x]'[(I-P\{s\})x])^{-1} [(I-P\{s\})x]' \qquad \text{(A17)}$$

Equations (A16) and (A17) can be used to find a recursion for $V'[I-P\{s+x\}]W$ in terms of $P\{s\}$, x, V, and W, all of which are known. Pre-multiplying (A17) by V' and post-multiplying by W yields

$$\begin{aligned}
V'[I-P\{s+x\}]W &= V'(I-P\{s\})W \ - \\
&\qquad V'(I-P\{s\})x[x'(I-P\{s\})x]^{-1} x'[(I-P\{s\}]W \\
&= [V'(I-P\{s\})V]^{1/2}[V'(I-P\{s\})V]^{-1/2} V(I-P\{s\})W \\
&\quad [W(I-P\{s\})W]^{-1/2} [W'(I-P\{s\})W]^{1/2} - V'(I-P\{s\})x \\
&\qquad [x'(I-P\{s\})x]^{-1/2} [x'(I-P\{s\})x]^{-1/2} x'[(I-P\{s\}]W \\
&= [V'(I-P\{s\})V]^{1/2} r\{s\}(V,W) [W'(I-P\{s\})W]^{1/2} - \\
&\quad [V'(I-P\{s\})V]^{1/2}[V'(I-P\{s\})V]^{-1/2} V'(I-P\{s\})x \\
&\quad [x'(I-P\{s\})x]^{-1/2} [x'(I-P\{s\})v]^{-1/2} x'[(I-P\{s\}]W \ * \\
&\qquad [W'(I-P\{s\})W]^{-1/2} [W'(I-P\{s\})W]^{1/2}
\end{aligned}$$

= [V'(I-P{s})V]^1/2 [r{s}(V,W) - r{s}(V,x) r{s}(x,W)] *

[W'(I-P{s})W]^1/2 (A18)

The resulting recursion relation is

V'[I-P{s+x}]W = [V'(I-P{s})V]^1/2

[r{s}(V,W)-r{s}(V,x) r{s}(x,W)][W'(I-P{s})W]^1/2 (A19)

Evaluating this equation at W = V and V = W gives

V'[I-P{s+x}]V^1/2 = [V'(I-P{s})V]^1/2 [I - r{s}(V,x)^2]^1/2

W'[I-P{s+x}]W^1/2 = [W'(I-P{s})W]^1/2 [I - r{s}(W,x)^2]^1/2

(A20)

Equation (A4) can be rewritten for projection onto {s+x}

r{s+x}(V,W) = [V'(I-P{s+x})V]^-1/2 V'[I-P{s+x}]W *

[W'(I-P{s+x})W]^-1/2

= [V'(I-P{s})V]^-1/2 [I - r{s}(V,x)^2]^-1/2 *

[V'(I-P{s})V]^1/2 [r{s}(V,W) - r{s}(V,x) r{s}(x,W)] *

[W'(I-P{s})W]^1/2 [W'(I-P{s})W]^-1/2 [I - r{s}(W,x)^2]^-1/2

r{s+x}(V,W) = [I - r{s}(V,x)^2]^-1/2 [r{s}(V,W) -

r{s}(V,x) r{s}(x,W)] [I - r{s}(W,x)^2]^-1/2 (A21)

This recursion can be used to find an order update for e~(p,t). Using the definition of e~(p,t) in equation (A5)

e~(p+1,t) = r{Yp+1(t)}(h,y(0:t)) (A22)

the matrix update can be applied when the projection space is changed from Yp(t) to Yp+1(t). To use equation (A21) let s = Yp(t) and let x = y(0:t)p+1. The space spanned by s and x is the same as the space spanned by Yp+1(t). Using the update recursion (A21) where $V = h$ and $W = y(0{:}t)$

$$\tilde{e}(p+1,t) = [1 - r\{Y_p(t)\}(h,y(0{:}t)_{p+1})^2]^{-1/2} * \\ [r\{Y_p(t)\}(h,y(0{:}t)) - r\{Y_p(t)\}(h,y(0{:}t)_{p+1}) \\ r\{Y_p(t)\}(y(0{:}t)_{p+1},y(0{:}t))] * \\ [1 - r\{Y_p(t)\}(y(0{:}t),y(0{:}t)_{p+1})^2]^{-1/2} \qquad \text{(A23)}$$

A similar regression can be found for the backward prediction error. The p+1 order predictor equations are

$$\begin{bmatrix} 0 \\ r(p+1,0) \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ r(p+1,t) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \\ y(0) \\ \vdots \\ y(t-p-1) \end{bmatrix} + \begin{bmatrix} 0 & & \cdots & 0 \\ y(0) & 0 & \cdots & \vdots \\ y(1) & y(1) & 0 & \vdots \\ & & & \vdots \\ y(p+1) & & . & y(1) \\ & & & \\ y(t) & & & y(t-p) \end{bmatrix} \begin{bmatrix} b(p+1,p+1) \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ b(1,p+1) \end{bmatrix} \qquad \text{(A24)}$$

The space spanned by $Y_{p+1}(t+1)$ is the same as the space spanned by $Y_p(t)$ and $y(0{:}t)$. The normalized backward prediction error $\tilde{r}(p+1,t)$ is defined as

$$\tilde{r}(p+1,t) = r\{Y_{p+1},t+1\}(h,y(0{:}t)_{p+1}) \qquad \text{(A25)}$$

Using the update recursion (A21) where $x = y(0{:}t)$, $s = Y_p(t)$, $V = h$, and $W = y(0{:}t)_{p+1}$

$$\tilde{r}(p+1,t) = [1 - r\{Y_p(t)\}(h,y(0{:}t))^2]^{-1/2} * \\ [r\{Y_p(t)\}(h,y(0{:}t)_{p+1}) - \\ r\{Y_p(t)\}(h,y(0{:}t))r\{Y_p(t)\}(y(0{:}t),y(0{:}t)_{p+1})] * \\ [1 - r\{Y_p(t)\}(y(0{:}t)_{p+1},y(0{:}t))^2]^{-1/2} \qquad \text{(A26)}$$

The recursion for the reflection coefficient $k(p,t)$ will be a time update. That is, $k(p,t)$ will be found from $k(p,t-1)$. The update follows from the defintion of the reflection coefficient

$$k(p+1,t) = L(p+1,t)/\mathrm{sqrt}(Rr(p)Re(p)) \tag{A27}$$

$$= [0\ r(p,0)\ldots r(p,t-1)] \begin{bmatrix} e(p,0) \\ \vdots \\ e(p,t) \end{bmatrix} 1/\mathrm{sqrt}(Rr(p)Re(p))$$

$$= [0\ r(p,0)\ldots r(p,t-2)] \begin{bmatrix} e(p,0) \\ \vdots \\ e(p,t-1) \end{bmatrix} 1/\mathrm{sqrt}(Rr(p)Re(p))$$

$$+\ r(p,t-1)e(p,t)/\mathrm{sqrt}(Rr(p)Re(p))$$

$$= [0\ r(p,0)\ldots r(p,t-2)] \begin{bmatrix} e(p,0) \\ \vdots \\ e(p,t-1) \end{bmatrix} 1/\mathrm{sqrt}(Rr(p)Re(p))$$

$$+\ r\tilde{}(p,t-1)e\tilde{}(p,t) \tag{A28}$$

Noting that $Re(p,t-1) + e(p,t)\hat{}2 = Re(p,t)$

$$Re(p,t-1) = Re(p,t) - e(p,t)\hat{}2$$

$$1/Re(p,t) = (1 - e\tilde{}(p,t)\hat{}2\ )/Re(p,t-1)$$

$$1/\mathrm{sqrt}(Re(p,t)) =$$

$$(1 - e\tilde{}(p,t)\hat{}2\ )\hat{}1/2\ /\mathrm{sqrt}(Re(p,t-1)) \tag{A29}$$

Inserting (A29) into (A28),

$$k(p+1,t)=[0\ r(p,0)\ldots r(p,t-2)] \begin{bmatrix} e(p,0) \\ \vdots \\ e(p,t-1) \end{bmatrix} (1 - e\tilde{}(p,t)\hat{}2\ )\hat{}1/2*$$

$$(1-r\tilde{}(p,t-1)\hat{}2)\hat{}1/2/\mathrm{sqrt}(Re(p,t-1)Rr(p,t-1))$$

$$+\ r\tilde{}(p,t-1)e\tilde{}(p,t)$$

$$= (1-e\tilde{}(p,t)\hat{}2)\hat{}1/2\ k(p+1,t-1)(1 - r\tilde{}(p,t-1)\hat{}2\ )\hat{}1/2$$

$$+\ r\tilde{}(p,t-1)e\tilde{}(p,t) \tag{A30}$$

The recursions necessary to implement the lattice filter are now available. The structure becomes more apparent when the recursions for $\tilde{r}(p,t)$ and $\tilde{e}(p,t)$ are written

$$\tilde{e}(p,t) = r\{Yp(t)\}(h,y(0:t)) \tag{A31}$$

$$\tilde{r}(p,t) = r\{Yp(t)\}(h,y(0:t)p+1) \tag{A32}$$

Using the order update recursions

$$\tilde{e}(p+1,t) = [I - r\{Yp(t)\}(h,y(0:t)p+1)^2]^{-1/2} *$$

$$[r\{Yp(t)\}(h,y(0:t)) - r\{Yp(t)\}(h,y(0:t)p+1) *$$

$$r\{Yp(t)\}(y(0:t)p+1,y(0:t))] *$$

$$[I - r\{Yp(t)\}(y(0:t),y(0:t)p+1)^2]^{-1/2}$$

$$\tilde{e}(p+1,t) = [1-\tilde{r}(p,t-1)^2]^{-1/2}[\tilde{e}(p,t)-k(p+1,t)\tilde{r}(p,t-1)]*$$

$$[1-k(p+1,t)^2]^{-1/2} \tag{A33}$$

Similarly, for $\tilde{r}(p+1,t)$,

$$\tilde{r}(p+1,t) = [I - r\{Yp(t)\}(h,y(0:t))^2]^{-1/2} *$$

$$[r\{Yp(t)\}(h,y(0:t)p+1) - r\{Yp(t)\}(h,y(0:t)) *$$

$$r\{Yp(t)\}(y(0:t),y(0:t)p+1)] *$$

$$[I - r\{Yp(t)\}(y(0:t)p+1,y(0:t))^2]^{-1/2}$$

$$\tilde{r}(p+1,t) = [1-\tilde{e}(p,t)^2]^{-1/2} [\tilde{r}(p,t-1)-\tilde{e}(p,t)k(p+1,t)] *$$

$$[1-k(p+1,t)^2]^{-1/2} \tag{A34}$$

LIST OF REFERENCES

LIST OF REFERENCES

[1] G. Hoshal, M. Siegel and R. Zapp, "A Microwave Heart Monitor and Life Detection System," Proceedings - Sixth Annual Conference I.E.E.E. Engineering in Medicine and Biology Society, Los Angeles, pp. 331-333, September 1982

[2] M.A. Popovic, K. Chan and J.C. Lin, "Microprocessor-based Noncontact Heartrate/ Respiration Monitor," ibid., pp. 754-755

[3] M. Nowogrodzki, D.D. Mawhinney and H.F. Milgazo, "Non-Invasive Instruments for the Measurement of Respiration and Heart Rates," I.E.E.E. 1984 National Aerospace and Electronics Conference Proceedings, Dayton, pp. 958-960, May 1984

[4] D.T.L Lee and M. Morf, "A Novel Innovations Based Pitch Detector," 1980 Proc. I.E.E.E. A.S.S.P., Denver, pp. 40-44, June 1980

[5] S.M. Kay and S.L. Marple, Jr. "Spectrum Analysis - A Modern Perspective," I.E.E.E. Proceedings, Vol. 69, No. 11, Nov. 1981,pp. 1380-1419

[6] ibid., pg. 1394

[7] A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw-Hill, New York, 1984 pg. 407

[8] ibid., pg. 409

[9] ibid., pg. 430

[10] J. Makhoul, "Linear Prediction: A Tutorial Review," I.E.E.E. Proceedings, Vol. 63, pp. 561-581, April 1975

[11] ibid., pg. 566

[12] G.E.P. Box and G.M Jenkins, Time Series Analysis, Forecasting and Control, Holden-Day Inc., San Fransisco, 1976, pg. 497

[13] G.W.J. Granger and P. Newbold, Forecasting Economic Timeseries, Academic Press, Orlando, 1977, pg. 72

[14] B Widrow, J.M. McCool, M.G. Larimore, C.R. Johnson,Jr. "Stationary and Non-stationary Learning Characterisics of the LMS Adaptive Filter," I.E.E.E. Proceedings, Vol. 64, No. 8, August 1976, pp. 1151-1162

[15] L.J. Griffiths,"A Continuously Adaptive Filter Implemented as a Lattice Structure," Proc. I.E.E.E. Conference on Accoustics, Speech and Signal Processing, Washington D.C., May 1977, pp.1151-1162

[16] B. Friedlander, "Lattice Filters for Adaptive Processing," I.E.E.E. Proceedings, Vol. 70,pp. 829-867, August 1982

[17] S.J. Orfanidis, Optimum Signal Processing, Macmillan, N.Y., 1984, Ch. 6, pg. 44

[18] ibid., pg. 313

[19] ibid., pg 312

[20] A. Papoulis,"Maximum Entropy and Spectral Estimation: A Review," I.E.E.E. Transactions on Accoustics, Speech, and Signal Processing, Vol. ASSP-29, No. 6, December 1981, pp. 1176-1186

[21] D.T.L. Lee, M. Morf, and B. Friedlander,"Recursive Least Sqaures Ladder Estimation Algorithms," I.E.E.E. Transactions on Accoustics, Speech, and Signal Processing, Vol. ASSP-29, No. 3, June 1981, pp. 627-641

[22] G.C. Goodwin and K.S. Sin, Adaptive Filtering, Prediction, and Control, Prentice Hall, Englewood Cliffs, 1984, pg. 58

[23] ibid., pg. 314

[24] ibid., pg. 329

[25] ibid., pg. 64

[26] A. Nehorai and M. Morf, "Enhancement of Sinusoids in Colored Noise and the Whitening Performance of Exact Least Squares Predictors," I.E.E.E. Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-30, No. 3, June 1982, pp. 353-362

[27] E.A. Robinson and M.T. Silvia, Digital Signal Processing and Time Series Analysis, Holden-Day, 1978, pg. 363

[28] V. Barnet and T. Lewis, Outliers in Statistical Data, John Wiley and Sons, Chichester, 1978, pg. 90

[29] C.R. Johnson, Jr., "Adaptive IIR Filtering: Current Results and Open Issues," I.E.E.E. Transactions on Information Theory, Vol. IT-30, No. 2, March 1984, pp. 237-250

[30] S.T. Alexander, E.H. Satorious, and J.R. Zeidler, "Linear Prediction and Maximum Entropy Spectral Analysis of Finite Bandwidth Signals in Noise," Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, 1978, pp. 188-191

[31] D.V. Bhaskar Rao and Sun-Yuan Kung, "Adaptive Notch Filtering for the Retrieval of Sinusoids in Noise," I.E.E.E. Transactions on Accoustics, Speech, and Signal Processing, Vol. ASSP-32, No. 4, August 1984, pp. 791-802