Linear Signal Processing and Ultrasonic Transversal Filters

WILLIAM D. SQUIRE, HARPER J. WHITEHOUSE, MEMBER, IEEE, AND JAMES M. ALSUP, MEMBER, IEEE

Invited Paper

Abstract—The role of linear transversal filters in signal processing is discussed in Section I. Linear filters for signal processing must often have complicated impulse responses, with large bandwidth and large time-bandwidth product. The linear transversal filter, a delay line with weighted and summed taps, is ideally suited for the implementation of such filters because of its simplicity of synthesis. The filter’s impulse response is derived by the application of some concepts from the theories of vector spaces and sampling, and is shown to be equal to the tap weighting function. Thus, the synthesis procedure merely consists of sampling the specified impulse response at appropriate intervals and using the sample values as the tap weights.

The utility of the transversal filter in signal processing is illustrated by an example from scatterer distribution mapping. The illustration is applied to two hypothetical systems—a sonar and an astronomical radar. In both these cases, it is not possible for a single filter to perform the signal in real time. Signal processing in compressed time is discussed as an alternative to the use of a large number of filters in parallel. If the processing filter has a bandwidth capability in excess of the signal’s bandwidth, the signal can be time compressed and processed serially in time. A generalized receiver, employing time compression, frequency translation, and multiple-input-port transversal filtering, is developed from these ideas.

In Section II, a generalized transversal filter is described and analyzed. A delay line with multiple arrays of taps, each array with a multiplicity of weighting functions, has as the impulse response between any pair of ports the cross-correlation function of the weighting functions for the two ports.

A number of implementations of transversal filters employing a variety of delay line types are described and some aspects of transduction and wave propagation in bounded media are presented in relation to these implementations.

INTRODUCTION

LINEAR systems play a leading role in science and engineering. In particular, linear filters, which occupy a position of universal importance in the electrical, electronic, and communication sciences have been studied extensively. There are three major classes of linear filters. The earliest developed was the “longitudinal” filter, in which the signal is operated upon successively by the structural elements of the filter. The ladder network and the meander line are typical examples. Because of its importance, and since it is the earliest of the three types, the longitudinal filter is in an advanced state of development. The second class is the “recursive” filter, in which previous values of the signal are modified and recombined with new values through feedback loops. The linear feedback network is a typical example of the recursive filter, which also has been developed to an advanced state.

The third and newest class is the “transversal” filter, in which the signal propagates in a delay line with transverse taps distributed along its length; the output is formed by weighting and summing the contributions from the taps (Fig. 1). The transversal filter, originally conceived for the
equalization of television signals, has been used extensively as an equalizing filter in telephone systems. In these applications, electromagnetic delay lines, both distributed and lumped, have been employed, but the total delay time available in filters with practical physical dimensions has been limited by the large propagation velocity of electromagnetic waves.

The use of acoustic delay lines results in a reduction in size, since the propagation velocity for acoustic waves in solids is $10^4$ to $10^5$ times slower than for electromagnetic waves. With the variety of methods now available for tapping microwave acoustic devices, the tapped ultrasonic delay line provides a practical means for implementing transversal filters. As a consequence, the transversal filter has many interesting and useful applications in microwave technology, as well as in filtering technology in general.

This paper presents the theory and properties of transversal filters, along with selected ultrasonic delay line implementations. It will be shown that the transversal filter, especially as embodied in contiguously tapped ultrasonic delay lines, is an ideal means of implementing linear signal processors. The paper is in two sections. In the first section, a time-space equivalence, based on some concepts from vector spaces and sampling theory, is applied to transversal filter analysis and synthesis, and the use of the transversal filter in a particular application—the mapping of a distribution of scatterers—is described in detail; the second section presents the theory of ultrasonic delay-line transversal filters and describes some implementations by means of magnetostrictive, piezoelectric, and optoacoustic taps.

I. LINEAR SIGNAL PROCESSING

Physical signals can be represented as vector elements in the Hilbert space of square integrable functions. A linear filter is a linear operator over this space and its impulse response is the kernel of the operator. If the filter is invariant under translation, its output is the convolution of the input signal with the impulse response, which is also an element of the space. The output also can be represented at each point as the inner product of the signal vector with a reversed and shifted replica of the impulse response vector [1].

A signal can be assumed without loss of generality to have finite bandwidth in any physical system, and can be represented by samples taken at intervals approximately equal to the reciprocal of this bandwidth. The sequence of these samples is the sampled signal; the original signal is reconstructed from it by appropriate linear filtering [2]. The space of sampled signals is a linear vector space whose vectors have the sample values as components [1]. By appropriately defining the signal's bandwidth $W$ and its length $T$, it is possible to approximate the sampled signal vector by a finite number $TW$ of components [3].

All of the preceding remarks about sampled signals apply equally to signals in time and in space [4]. The taps on a delay line perform a spatial sampling of the signal in the delay line and if the taps are ideal (i.e., if their impulse responses are $\delta$ functions), the sequence of tap outputs is the spatially sampled signal. If the tap outputs are weighted and summed, the summation yields a linear operator whose kernel is the weighting function. The sum can be regarded as the output of either a space-domain or a time-domain filter. This kind of linear filter is called a transversal filter. Kallmann introduced the transversal filter in 1940, for the special case of a lossless, nondispersive delay line with lightly coupled, nonreflecting taps [5]. Since the signals in the delay line, at the tap outputs, and at the output of the summing circuit are time varying, they can be filtered in the time domain as well as in the space domain.

The tap weights $w_i$ can be regarded as the sample values $w(\tau_i)$ of a weighting function $w(t)$, where $\tau_i$ is the delay to the $i$th tap. These weights can also be regarded as the sample values $v(\xi_i)$ of a corresponding space weighting function $v(\xi) = w(\xi/c)$, where $\xi$ is the distance to the $i$th tap and $c$ is the propagation velocity of waves on the line. It is assumed that all taps have the same impulse response $u(t)$. Thus, the time-varying function that appears at the output of the $i$th tap when an impulse is applied to the line is $u(t)$ weighted by $w_i$ and delayed by $\tau_i$. The impulse response of the filter $h(t)$ is the summation of these individual impulse responses:

$$h(t) = \sum_i w_i u(t - \tau_i). \quad (1)$$

For all of the implementations of transversal filters to be discussed in this paper, the taps have a bandpass filter characteristic, and the reconstruction of $w(t)$ from the sample values $w_i$, is accomplished by the filtering action of the taps, or by this action combined with that of a simple compensating filter at the input or output of the transversal filter. If $u(t)$ is redefined to include this compensation, when required, then the right hand side of (1) is identical in form to the expression representing reconstruction of the function $w(t)$ from the samples $w_i$[2], and (1) reduces to $h(t) = w(t)$.

The output of a transversal filter for a general input signal is then the convolution of that signal with the filter's tap-weighting function:

$$y(t) = [x * w](t) \quad (2)$$

where $y(t)$ is the output signal, $x(t)$ is the input signal, and $*$ denotes convolution.

Thus, transversal filter synthesis, in terms of its time-domain specification $h(t)$, is extremely simple. The synthesis procedure consists merely of sampling the specified $h(t)$ at intervals corresponding to its bandwidth, using these sample values as weights on appropriately spaced delay-line taps, and summing the outputs of these weighted taps.

![Transversal Filter Diagram](image-url)
Scatterer Distribution Mapping

As an example of the applications of transversal filters to linear signal processing, the mapping of scatterers distributed in range and Doppler will be described.1 Such a mapping is of interest in sonar [6], radar [7], and radar astronomy [8]. A finite-energy signal \( s(t) \) is transmitted and the scattered energy is received in the presence of white Gaussian noise. For the detection of \( s(t) \) in such noise, both a matched filter and a correlator are optimum [9]. A matched filter (Fig. 2) is a linear filter whose impulse response \( h(t) \) is the time-reversed replica of the signal to be detected: \( h(t) = s(-t) \). A correlator (Fig. 3) multiplies the input, consisting of signal and noise, by a locally formed replica of the signal and integrates the product.

The mapping of a distribution of scatterers is a more complex problem than that of simple signal detection; in this case, the matched filter and correlator are not always optimum. Nevertheless, even when they are not optimum, they often provide acceptable performance and when they do not, slightly mismatched versions may provide nearly optimum performance [10]. As a consequence, these processors are used extensively.

The mapping of a distribution of scatterers by either of these processors can be expressed in terms of the resolution function \( \chi(x, v) \). This function represents the response of the processor when matched to \( s(t) \) for an input, \( s_r, o(t) \), which is a version of \( s(t) \) delayed by \( \tau \), with Doppler \( \nu \), where \( \nu = 2(R/c)f_0 \), \( R \) is the range rate, \( c \) is the propagation velocity, and \( f_0 \) is the mean frequency.2 Since \( \chi(x, v) \) is the response of a processor matched to \( s(t) = s_0, o(t) \), it is the convolution of \( s_r, o(t) \) with \( s_0, o^- \):

\[
\chi(x, v) = s_r, o(t) * s_0, o^-(3)
\]

where \( s_r, o^-(t) = s_r, o(-t) \).

To map the point \( \tau_0, v_0 \) in the range-Doppler space, the processor is matched not to \( s(t) \) but to \( s_{r, o}(t) \). However, the processor when so matched responds not only to the signal scattered from the point \( \tau_0, v_0 \), but also to scattering from the other points in the range-Doppler space. The output of the processor for scattering from the point \( \tau, v \) when matched to the point \( \tau_0, v_0 \), is \( \chi(\tau, v) = s_r, o * s_{r, o}^\tau \). \( \chi(\tau, v) \) has been translated from the origin to the point \( \tau_0, v_0 \) has suffered a slight distortion in shape in the process of translation. If \( \chi(\tau, v) \) is concentrated in the region of \( \tau_0, v_0 \) or equivalently if \( \chi \) is concentrated at the origin, then the processor output will be a good estimate of the scattering cross section, averaged over the region of concentration. Thus the size of the region around the origin in which \( \chi \) is concentrated determines the resolution with which the range-Doppler space is mapped [11]–[13].

The nominal resolution for a signal with bandwidth \( W \) and duration \( T \) is \( 1/W \) in \( \tau \) and \( 1/T \) in \( v \). The map of the scattering distribution that results from such processing is a function of \( \tau \) and \( v \) and can be regarded as having bandwidth in both these variables; the nominal bandwidth in \( \tau \) is \( W \) and in \( v \) is \( T \). A straightforward extension of one-dimensional sampling theory to two dimensions leads to the result that \( TW \) samples per unit square in the range-Doppler space (i.e., one sample for each resolution cell of size \( 1/W \) in \( \tau \) and \( 1/T \) in \( v \)) span the entire space. The continuous map of the scattering distribution can be recovered from these samples by appropriate smoothing.

Although the nominal resolution in \( \tau \) and \( v \) is \( 1/W \) and \( 1/T \) for any signal with bandwidth \( W \) and duration \( T \), the theory of signal design leads to many choices of signal form, with corresponding variations in the form of \( \chi(\tau, v) \). A selection from among these choices must be made on the basis of prior knowledge of the scattering distribution as well as the desired mapping resolution. Signal design is an active area of current research [14]. The mapping may require any signal with time-bandwidth product \( TW \), if the resolution in the range-Doppler space is to be \( \delta_\tau = 1/W \) and \( \delta_v = 1/T \). If the filter synthesis procedure is to be sufficiently versatile to provide a processor with the required resolution for the mapping of any scattering distribution encountered, then this procedure must be capable of synthesizing a filter to match any signal with this time-bandwidth product. When the \( TW \) product becomes large, the frequency domain synthesis of conventional longitudinal filters matched to arbitrary signals specified in the time domain becomes increasingly difficult. On the other hand, the synthesis of transversal filters remains simple, since the procedure consists merely of selecting the tap weights to correspond to the sample values of the impulse response.

Suppose it is desired to map a distribution of scatterers in the range interval \( \Delta R = R_{max} - R_{min} \), and the range-Doppler interval \( \Delta v = v_{max} - v_{min} \), with a prescribed resolution in range \( \delta R \) and range-Doppler rate \( \delta \nu \). Since \( \delta R = c\delta /2 \) and \( \delta \nu = \nu \delta /2 \), the required signal duration and bandwidth are

\[
T = \lambda_0 / 2 \delta R
\]

and

\[
W = c/2 \delta \nu
\]

where \( \lambda_0 \) is the wavelength corresponding to the mean fre-
From the second equation of (4), a relation can be established between the system’s fractional bandwidth and the number of range resolution cells per unit wavelength:

\[ W/f_0 = \frac{\lambda_0}{2\delta_R} \]  

(5)

where \( W/f_0 \) is the fractional bandwidth. From (4), the required signal-time bandwidth product is given by

\[ TW = \frac{c\lambda_0}{4\delta_R \delta_k} \cdot \]  

(6)

The time required to form the receiver output for a given range \( R \) and range rate \( \dot{R} \), within a resolution cell of size \( \delta_R \times \delta_k \) is the resolution cell processing time \( t_p \). The duration of the reflection from scatterers within the range interval \( \Delta_R \) is \( (2\Delta_R/c) + T \), and the effective resolution-cell processing time for a matched filter, \( t_{p[mf]} \), is this duration divided by the number of range resolution cells, \( \Delta_R/\delta_R \), or

\[ t_{p[mf]} = \frac{[2\Delta_R/c] + T}{\delta_R/\Delta_R} \]  

However, for the correlator, the resolution-cell processing time is simply the signal duration

\[ t_{p[cor]} = T \]  

The time between transmissions is the pulse interval \( t_i \), and is available for signal processing, so that the maximum number of computations possible between transmissions is

\[ N_p = t_i/t_p \]  

and depends on whether a matched filter or a correlator is being used.

\[ N_{p[mf]} = t_i/[2\Delta_R/c + (T\delta_R/\Delta_R)] \]  

(7a)

\[ N_{p[cor]} = t_i/T. \]  

(7b)

The number of computations \( N \) required to span an area \( \Delta_R \times \Delta_k \) in the range-range-rate space is the product of the number of range resolution cells \( \Delta_R/\delta_R \) and the number of range-rate resolution cells \( \Delta_k/\delta_k \): \( N = (\Delta_R/\delta_R)(\Delta_k/\delta_k) \). The required number of computations can be executed during the interval \( t_i \) only if \( N_p/N \geq 1 \).

Since \( N \) is the same for both the matched filter and the correlator, the ratio \( N_{p[mf]}/N_{p[cor]} \) is a measure of the processing speed of the matched filter relative to the correlator. From (7a) and (7b),

\[ \frac{N_{p[mf]}}{N_{p[cor]}} = \frac{\alpha TW}{N} \]  

(8)

where \( \alpha = 1/[1 + (cT/2\Delta_R)] \) and \( 0 < \alpha < 1 \). Only when \( \alpha < 1/TW \), does the correlator have greater processing speed than the matched filter.

These ideas will now be illustrated with numerical examples from sonar and radar astronomy. Although the examples are not intended to correspond to actual systems, they are representative of the parameters that might be required for ocean-floor or planetary mapping.\(^9\) The system parameters, along with the signal parameters required to achieve these system parameters, are presented in Table I, and the corresponding processor performance parameters are presented in Table II. Note that in the sonar example, \( \Delta_R \) is not small relative to \( c/2TW \) and, therefore, the frequency shift approximation for Doppler is not valid over the entire range-rate interval. However, the nominal range-rate resolution is still \( 1/T \). A method for treating this case by dividing

\(^9\) Since both the general expressions and the numerical values are shown, other systems can be compared either by direct calculation or by scaling.
Δk into subintervals which are small relative to c/2TW is discussed later.

From Table I, the bandwidth and TW product for the sonar example are 750 Hz and 750, and for the radar astronomy example are 6 kHz and 1.5×10⁴, respectively. Ultrasonic transversal filters have been constructed with bandwidths to 1 MHz and TW products to 10⁴, and filters are under development with bandwidths to 1 GHz and TW products of 10⁴ [15], [16]. Thus, existing filters have bandwidths in excess of the requirements of sonar and radars discussed later. Since the TW product is a measure of the maximum possible mapping resolution, sonar resolution could be increased by more than an order of magnitude beyond the value assumed in this example without exceeding the TW capacity of transversal filters currently under development. Conversely, even these latter filters have TW capacities which are an order of magnitude smaller than that required for the radar astronomy example. It is hoped that advances in the technology of microwave acoustic delay lines will provide the increase in TW capacity needed for such radar astronomy applications.

From Table II, the speed with which the matched filter can map the distribution of scatterers is approximately 400 times that of the correlator, for both the sonar and radar astronomy examples. However, the values of Np/NT presented in Table II indicate that even the matched filter is approximately 100 times too slow to accomplish the mapping in real time, for either the sonar or the radar.

The simultaneous use of a sufficient number of individual filters, each matched to a different Doppler modification of the signal, will permit real-time processing. An alternative method that uses a single filter and employs time compression and serial processing will now be described. If desired, a combination of these techniques may be used.

If the signal to be processed is compressed in time by the factor k, then the resolution cell processing time t_q is reduced by this factor and the number of computations N_p, during the interval t_i, is increased by this factor. Thus, time compression by the factor k permits a k-fold increase in processing speed. For the examples given earlier, real-time processing can be accomplished with a single matched filter if k equals 77 for the sonar or 140 for the radar.

The bandwidth of a time-compressed signal is increased by the factor k and the implementation of a matched filter for a time-compressed signal must have a bandwidth capability k times larger than the bandwidth of the corresponding real-time signal. If, by a particular implementation, a transversal filter can be constructed which is capable of producing impulse responses with a maximum bandwidth of W_max and a maximum duration of T_max, the filter can be used for time-compressed processing of any signal with bandwidth W and duration T such that W < W_max and TW ≤ T_max W_max; the maximum compression factor is k_max = W_max/W. Thus, bandwidth capability in excess of the bandwidth of the real-time signal makes possible the use of time compression for the sonar and radar astronomy example, the excess bandwidth of existing filters makes possible values of k large enough to permit real-time processing by a single filter.

Since k_max is proportional to W_max, then W_max is a figure of merit for the filter's processing speed, and since the maximum possible number of resolution cells per unit cell is proportional to the signal's time-bandwidth product TW, then T_max W_max is a figure of merit for the filter's resolution capabilities. Using current ultrasonic technology, filters and time compressors have been built which, for most sonar systems, have a k_max greater than 10⁴ [17], and microwave acoustic devices are under development that will make possible values of k_max greater than 10⁴ [18]. However, in many radar systems the signal bandwidth limits k_max to small values for these same filters. It is hoped that additional developments in microwave acoustics will soon make larger time compressions possible.

A Generalized Linear Receiver

The use of time compression with matched filtering is illustrated by the linear receiver shown in Fig. 4. The matched filter shown in Fig. 4 has multiple-output ports, whose significance will be discussed later. A range-delayed and Doppler-modified signal s_s(t) is to be processed by the receiver whose bandwidth W_r must be wide enough to accommodate the possible Doppler shifts of the signal's spectral components. By means of single-sideband (SSB) modulation, the receiver shifts the spectrum of the input signal downward by the amount f_d = f_o - W_r/2, from the receiver's passband to its baseband, as shown in Fig. 5(a). The receiver's input is so shifted because the required sampling rate is often lower for baseband signals than for bandpass signals [2]. The lowest possible sampling rate is desirable, since the time compressor chosen for illustration is a sampled-data device for which the largest compression factor k_max is obtained when the sample rate is lowest. If the spectrum of the original signal is S(f), then the spectrum of the Doppler-modified version is S(βf), where $\beta = 1 + 2R/c$. It follows that the spectrum of the signal in the baseband is$h_s(f) = 0$ and the spectrum of the time-compressed baseband signal is $S((βf - f_s)/k)$, as shown in Fig. 5(b). A second SSB modulator shifts the time-compressed signal upward by the amount $f_s$ to the frequency band of the matched filter. The shifted spectrum is $S((βf - f_s)/k) + f_s$. The compression factor k has been chosen so that after time compression the signal bandwidth corresponds to the filter bandwidth $k = k_max$. If a single sample is taken from each cell of size $R_\theta \times 2R_\theta$ in the range-rate space, the resulting set of samples is sufficient for an approximate description of the scattering distribution. Since a matched filter processes only values of range for a single value of Doppler, this space can be mapped with filters matched to the values of $R$ at the

---

4 It will be shown later that excess bandwidth can be used to increase processing speed.

---

6 This shift is not required if the matched filter operates at baseband.
centers of strips of width $\delta_k$ parallel to the $R=0$ axis. Therefore, a total of $\Delta_k/\delta_k$ individual matched filters are required for real-time processing. When a filter operates in compressed time with a compression factor $k$, compressed blocks of signal can be processed serially at a $k$-fold greater rate. Thus, when $k \geq \Delta_k/\delta_k$, serial processing can be substituted for parallel processing and only a single filter is required.

Time compression and transversal filtering are the features that distinguish this receiver from a conventional superheterodyne receiver which employs only frequency translation and bandpass filtering. Frequency translation, usually accomplished by SSB modulation, changes the fractional bandwidth of the input while leaving its duration unchanged, whereas time compression changes the duration while leaving the fractional bandwidth unchanged. Thus, this generalized receiver provides complete flexibility in utilizing the performance characteristics of the matched filter and makes possible the efficient application of transversal filters, implemented with microwave acoustic devices, to the processing of signals with spectra anywhere in the frequency range from seismic through radio to microwave frequencies.

**Time Compression**

Time compression is achieved by storing the signal in memory and retrieving it at a rate greater than the rate of storage [19]. Although time compression can be achieved with analog devices operating on continuous signals [20], [21], sampled data devices are usually employed. The principle of time compression can be illustrated by a transversal filter with $p$ uniformly weighted taps (Fig. 6) which are separated by time intervals of duration $\Delta r_s = \Delta r - \Delta r_t$, where $\Delta r_s$ is the sample interval in real time, $\Delta r$ is the sample interval in compressed time, and $p$ is an integer greater than or equal to $TW$. At the instant when the oldest sample in the line passes the last tap, it appears at the output of the filter. $\Delta r_s$ seconds later, the next-to-oldest sample passes the next-to-last tap and appears at the output of the filter, and samples continue to pass their corresponding taps and appear at the output until $(p-1)\Delta r_s$ seconds later, at which time the newest sample passes the first tap and appears at the output of the filter. Thus, $p$ samples have appeared at the output of the filter in the proper sequence, but with a separation of $\Delta r_s$ instead of $\Delta r$; they have been compressed in time with a compression factor $k = \Delta r_s/\Delta r$. A pause of duration $(k-p+1)\Delta r_s$ follows, during which the next-to-oldest sample propagates the remaining distance to the last tap. The oldest of the $p$ samples has passed out of the delay line, all samples have propagated a distance corresponding to $\Delta r_s$ along the line, and a new sample has been introduced into the line. The process just described now repeats. Thus, the output of the time compressor consists of blocks of $p$ compressed samples with a gap of duration $(k-p+1)\Delta r_s$.
between blocks. Each block contains all of the samples from the previous block except the oldest, with the samples shifted within the block by $\Delta r_e$ relative to the positions of the samples within the previous block, and with a new sample added. These blocks can be reconstructed to form blocks of continuous signal by the filtering action of the taps themselves, if the tap impulse response is properly selected. The description of the operation given here is valid only when

$$(1 - 1/p)\Delta r_e < \Delta r_t < \Delta r_e. \tag{9}$$

When this inequality is not satisfied, the samples do not pass their corresponding taps sequentially in time and the situation becomes complicated.

The compressed signal lies entirely within one or more of these blocks only when the input signal lies entirely within the delay line. Therefore if $p > TW$, each of $q$ adjacent blocks, where $q = p - TW + 1$, will contain the entire signal; all other blocks will contain only part of the signal or none at all. A filter matched to a particular Doppler modification of the signal is connected to the output of the time compressor. At the end of a time interval $(p - 1)\Delta r_e$, equal to the duration of one of the output blocks, the impulse response of the filter can be altered to match a different Doppler modification of the signal. After another $\Delta r_e$ seconds have elapsed, the filter can again be modified to match another Doppler modification. This process is repeated at intervals of $\Delta r_e$ until all required Dopplers have been examined by a corresponding matched filter. Since there are only $q$ blocks that contain the entire signal, at most $q$ modifications of the filter’s response can be made before it must be returned to its initial state. Otherwise, the filter response that matches the signal may not be formed during the time that the signal is present in the time-compressor output.

The time-compression factor $k$ acts as a fundamental limit on the number of Dopplers that can be searched for all range values. The utilization factor $q/k$ measures the degree to which a time-compressor’s performance approaches this fundamental limit. In the compressor used for illustration, $k$ must be greater than or equal to $p$. Otherwise, the sample blocks would overlap and fewer than $q$ Dopplers could be searched. The optimum situation exists when $p$ can be made equal to $k$, for in this case the gap between blocks reduces to $\Delta r_e$ and $q$ has the maximum value possible for a fixed value of $k$. Then $q/k = 1 - (TW - 1)/k$ and, as $k$ increases, the utilization factor approaches unity. Thus, for the time compressor illustrated here, both large $k$ and large $p$ are desirable.

For most sonar systems, values of $k$ greater than $TW$ are necessary if all required Dopplers are to be searched for all ranges, and time compressors with sufficiently large $k$ have been developed for such systems [19]. However, for many radar systems, the required time-compression factors are less than $TW$ and the time compressor chosen here for illustration is not appropriate. For such radars, and for sonars with $k < TW$, time compressors with tap intervals that do not satisfy (9) must be used or other equivalent devices must be employed.

It may not be necessary to alter the filter’s response $q$ times to search $q$ Dopplers. Within a Doppler interval that is small relative to $c/2TW$, the modification of the signal as a result of Doppler can be approximated by a frequency translation of the signal. If the Doppler interval $\Delta g$ satisfies this condition, the search in Doppler can be accomplished by shifting the frequency of the variable-frequency oscillator (VFO) of the second SSB modulator by amounts corresponding to $\delta \tilde{g}$ at time intervals $\Delta r_e$ instead of by altering the filter’s response. It is still true, of course, that only $q$ Dopplers can be examined. If $\Delta g$ is too large for the frequency shift approximation to be valid over the entire interval, it can be divided into subintervals in which the frequency-shift approximation is valid. Then the filter must be altered to match the Doppler modification of the signal for the Doppler value at the center of each of these subintervals; within the subintervals, however, the Doppler search can be achieved by frequency shifts of the VFO.

Although this description of time-compressed processing is based on the assumption that the $q$ different modifications of the signal are Doppler modifications, the generalized receiver described here is not so restricted. The $q$ possible modifications of the filter’s impulse response could also represent signal modifications such as those resulting from acceleration. Even more generally, they could represent any set of $q$ orthogonal signals, such as those corresponding to $q$ symbols of an alphabet for use in communication or $q$ words in a coded message.

Although any type of memory can be used for a time compressor, computer-type memories are usually employed; the signal is sampled at baseband, digitized to form a binary number, and stored [19]. The ultrasonic delay-line time compressor (DELTIC), introduced by Anderson in 1956 [22] for the compression of sampled signals digitized at two levels, employed a computer-type serial quartz delay-line memory, which was a forerunner of modern recirculating memory stores [23]. Configurations of the DELTIC exist that produce blocks of samples identical in form and compression factor to those produced by the transversal filter time compressor (TFTC), both for the simple case illustrated here and for the more complicated cases that occur when (9) is not satisfied. Thus, there is no difference between the performance of a DELTIC and a corresponding TFTC. However, the total delay time required in the DELTIC is approximately $\Delta r_e$, whereas in the TFTC it is $p\Delta r_e$. Note that $p\Delta r_e$ is greater than or equal to $T$, the signal’s duration, and the requirement that the TFTC have this much delay makes it impractical for many systems. It was introduced here for illustrative purposes because of its simplicity. Conversely, the DELTIC is practical for application to many systems because of its relative short delay-time requirements. Its delay-time advantage is obtained, however, at the expense of requiring binary digital signals at the input, since the recirculating delay-line amplitude stability is generally insufficient to preserve the sample amplitudes of analog or multilevel digital signals during the number of recirculations required.

To use the DELTIC for the compression of analog or multilevel digital signals, the samples must first be converted
to binary coded numbers, and then compressed either by an appropriate number of DELTICs operating in parallel, or by a single DELTIC operating serially at a correspondingly increased speed [24].

Multiple-Output-Port Transversal Filters

For many implementations of transversal filters, it is possible to realize a number of simultaneous outputs with a single delay line [25]. These multiple outputs are formed from an array of tap weights, as shown in Fig. 7. Multiple-output-port filters may provide savings in system size and complexity, and some techniques for their implementation are discussed in the second section of this paper.

Such a filter can replace a bank of filters operating in parallel to provide real-time processing. In addition, multiple-output-port matched filters are useful in increasing the effective speed of serial processing in compressed time. If such a filter operates with a time-compression factor \( k \) and has \( m \) output ports, and each port represents a different one of the impulse responses needed for the search in Doppler, then the filter has an effective processing speed of \( k \times m \). For a system with fixed Doppler search requirements, the addition of multiple output ports reduces the time compression required for Doppler search, and thus reduces the bandwidth capability required of the filter.

All of the earlier statements concerning the application of sampling in space and time to transversal filters apply also to multiple-output-port transversal filters. The tap weights \( w_{ij} \) can be regarded as the sample values \( w_j(t) \), of a weighting function \( w_j(t) \), where \( r_i \) is the delay to the \( i \)th tap and the subscript \( j \) indicates the port at which the output is observed. These weights can also be regarded as the sample values \( v_j(\xi_i) \) of a corresponding space weighting function \( v_j(\xi) = w_j(\xi/c) \), where \( \xi_i \) is the distance to the \( i \)th tap and \( c \) is the propagation velocity of waves on the line. The \( j \)th output has impulse response

\[
h_j(t) = \sum_i w_{ij} u(t - r_i) \tag{10}
\]

where \( w_{ij} \) is the weight of the \( i \)th tap for the \( j \)th output and \( u(t) \) is the tap impulse response. The filtering action of the taps, with compensation as required, reconstructs \( w_j(t) \) from the sample values \( w_{ij} \), and (10) reduces to \( h_j(t) = w_j(t) \). Thus, for a general input signal, the \( j \)th output of a multiple-output-port transversal filter is the convolution of that signal with the filter’s \( j \)th tap weighting function:

\[
y_j(t) = [x * w_j](t) \tag{11}
\]

where \( y_j(t) \) is the \( j \)th output signal and \( x(t) \) is the input signal. The synthesis of multiple-output-port transversal filters is comparable in simplicity to that for the single-output-port filter.

Conclusion

Because of the extreme simplicity of its synthesis, the transversal filter is ideal for applications, such as linear signal processing, where a complicated impulse response with a large time-bandwidth product may be required.

Existing ultrasonic delay-line transversal filters have many applications in sonar, radar, and radar astronomy systems. Microwave acoustic delay-line transversal filters under development can be expected to have many additional applications. Since generally the physical size of these devices decreases and the processing speed increases as the maximum bandwidth increases, it is to be hoped that the operation of acoustic delay-line transversal filters will be advanced still further into the microwave region.

In many systems, real-time signal processing is impractical because the number of simultaneous operations to be performed exceeds the system’s capability. If the bandwidth capability of the signal-processing filter exceeds the bandwidth of the signal to be processed, the signal can be time-compressed by an amount \( k \), proportional to the bandwidth ratio. When the input signal is thus compressed in time, the processing can proceed at a \( k \)-fold greater rate. Thus, linear filters operating in compressed time make possible significant increases in the effective speed with which linear signal processing can be performed, and transversal filters are ideal for this application because of the simplicity of the synthesis procedure required to obtain an arbitrary impulse response.

Some implementations, to be described in the second section of this paper, make possible the construction of transversal filters with multiple output ports, at each of which appears simultaneously the response to an impulse at the input port, each response corresponding to one of a set of mutually orthogonal functions. If the filter has \( m \) such ports, and processes at a \( k \)-fold increased rate, as in the generalized receiver described here, it is equivalent to \( k \times m \) single-output-port filters operating in parallel in real time.

It should not be inferred that the uses of transversal filters discussed so far are the only ones for which they are suited. The case with which time-domain synthesis of transversal filters with large \( TW \) product can be accomplished, and the availability of compact implementations employing ultrasonic delay lines, make transversal filters, in conjunction with time compression and frequency translation, useful in a wide range of applications encountered in diverse disciplines. Besides the many applications to scatterer-distribution mapping in the fields of radar, radar astronomy, and sonar, there are also mapping applications in seismology, nondestructive testing, and medical instrumentation. In addition, there are significant applications in interferometry, spectral analysis, communications, adaptive filtering, and delay-line technology, as well as in laboratory instrumentation.
A number of implementations of transversal filters and time compressors employing ultrasonic delay lines are currently available, some of which are discussed in the second section of this paper. Although these devices have made possible the solution of a number of important signal-processing problems, there are still many important problems for which acoustic transversal filters and time compressors with higher speed ($W$) and larger capacity ($TW$) are required. Because of the frequencies at which these needed devices must operate, they fall within the domain of microwave science and engineering. The special talents of the worker in microwaves might fruitfully be applied in advancing this technology.

II. ULTRASONIC TRANSVERSAL FILTERS

In the first section of this paper, the Kallmann transversal filter was described and some of its properties and applications were discussed. In this section, a generalized transversal filter is introduced and analyzed, some features of delay-line transduction and propagation are discussed in the context of transversal filter implementation, and a number of practical implementations employing magnetostrictive, piezoelectric, and optoacoustic taps on a variety of ultrasonic delay lines are described.

The Generalized Transversal Filter

In the Kallmann filter, it was assumed that the desired signal is propagating along the delay line where taps, which do not perturb the propagating wave, spatially sample, weight, and sum the signal. No consideration was given to the methods of transduction by which the wave was launched. There is often an advantage in using a transducer composed of a weighted array of elements. If its elements are reciprocal, the array can be used both for launching and tapping, which is a useful feature in some applications. The utility of filters with multiple-output ports was discussed in the first section of the paper, where a method employing a two-dimensional weighting function applied to a tap array was described. Another method uses a sequence of weighted arrays along the line. A natural extension is a filter with both multiple output ports and multiple-input ports. The most general filter would be composed of a sequence of nonoverlapping arrays along a delay line with absorbers at the ends of the line to prevent reflections, with a two-dimensional weighting function applied to each array location to form multiple ports, and with each array reciprocal and usable as both an input and an output array. Since several of the implementations of transversal filters described here have one or more of these features, the generalized transversal filter incorporating all of them is discussed. First, however, an application of such a filter with three sequential ports is described.

A filter whose impulse response is a time reversal of the transmitted signal is frequently used for detection or scatterer distribution mapping. The generation of the signal to be transmitted is straightforward when the signal is simple, but when it is complicated and has a large $TW$ product, sophisticated techniques of generation may be required. One method is to sample the signal and store the samples in memory, to be retrieved and reconstructed into the signal at each transmission. Another method is to generate the signal recursively by the application of an appropriate algorithm to an elemental signal. For signals with large $TW$ product, the former method requires a large memory, as does the latter method because of its complicated algorithm.

An attractive alternative is to use the same generalized transversal filter for both signal generation and signal processing. Consider a transversal filter with two transducers separated by a tap array. If $h(t)$ is the response at the output of the tap array when an impulse is applied to one of the transducers, then $h(-t)$ is the response when the impulse is applied to the other transducer. Thus, if $h(t)$ is the required signal for transmission, it can be obtained at each transmission merely by applying an impulse to the appropriate transducer; the other transducer is the input of a filter matched to $h(t)$. Such a three-port transversal filter is illustrated in Fig. 8. If the transducers and taps are reciprocal, an alternate method is to use the tap array as the input filter, with the transducers acting as outputs. If the response at one transducer due to an impulse applied to the tap array is $h(t)$, then the response at the other transducer is $h(-t)$, and again the filter can be used for both the generation and the matched filtering of $h(t)$. For the case of reciprocal transducers and taps, the impulse response between any pair of ports is independent of which port of the pair is used as the input. If the transducers are lightly coupled and nonreflecting, the impulse response between the transducers and the tap array is related to the tap weighting function in a manner analogous to the relation for the Kallmann filter: $h_{12}(t) = h_{12}(t) = w(t)$ and $h_{21}(t) = h_{21}(t) = w(-t)$, where the subscripts 1, 2, and 3 indicate the left end port, the right end port, and the tap array port, respectively (see Fig. 8), and where $w(t)$ is the tap weighting function expressed in the time domain.

An additional advantage of using the same filter for both signal generation and detection or mapping is that the constraints on variations in the total delay time of the line can be relaxed. When an independent signal generator is used, the total delay time of the transversal filter's delay line must be held constant to a small part of the delay time between taps; otherwise, as the total delay time is changed by temperature variations or other causes, the transversal filter will no longer match the transmitted signal. To achieve the necessary stability, isopaustic (constant delay) materials may be required. However, if the same transversal filter is used both for signal generation and detection, then as the total delay-time drifts, the matched filter tracks the generated signal. The only constraint on the total delay-time drift is
that it be small during the time interval between transmission
and reception of the echo or scattered signal when it is com-
pared with the delay time between taps.

Before proceeding with the analysis of the generalized
transversal filter, some aspects of transduction and tapping
must be examined. In the Kallmann filter, discrete distribu-
tions of point taps are weighted and summed to form the
output. In the generalized transversal filter, it is desirable
to consider continuous as well as discrete distributions of
elements both for launching and for tapping.

From the theory of sampling, any continuously distributed
array of elements can be represented as a discrete array,
with the element spacing approximately equal to \( c/W \), where \( c \) is the propagation velocity of waves on the line and \( W \) is
the maximum bandwidth of signals to be propagated along
the line. Conversely, a discrete array of elements with such
spacings and appropriate reconstruction filtering as is pro-
vided by properly compensated elements is equivalent to a
continuously distributed array with the same weighting func-
tion. When elements have this spacing, they are said to be
"contiguous." The element weights can be regarded as com-
ponents of a vector. This vector is the sample function, cor-
responding to a continuous weighting function.

The properties of an individual element of a discrete array
can be described in terms of its effective length. If an impulse
is applied to such an element, the resulting wave propagating
along the delay line is the element’s impulse response; the
length of this wave is the effective length of the element.
When an impulse wave propagating along the line passes
such an element, the resulting output is the element’s im-
pulse response; the effective element length is \( c \) times the
impulse response duration. If the element is reciprocal so
that it can be used both as a launch and as a receive element,
it has the same effective length in both modes. An element
whose effective length is substantially less than \( c/W \) can be
approximated by a point element with an impulse response
that is a delta function. If the effective length is comparable
to \( c/W \), the element can be represented as a point element
whose impulse response is the same as that of the actual
element. If the effective length is substantially greater than
\( c/W \), the element can be represented as an extended, dis-
tributed array.

The relation between the launch and receive array weight-
ing functions and the generalized transversal filter's impulse
response is obtained by an analysis parallelizing that given in
the first section of this paper for the Kallmann filter. The
analysis is developed for distributed launch arrays as well
as for distributed receive arrays. As in the case of the Kall-
mann filter, the delay line is assumed to be lossless and non-
dispersive and the array elements are assumed to be lightly
coupled and nonreflecting. The impulse response of the filter
is derived by the integration of continuously distributed
weighting functions. However, by regarding the resulting
expressions either as Stieltjes integrals or as integrals of
generalized functions, it can be shown that the expressions
apply to discrete as well as to continuously distributed
arrays.

Consider a particular continuously distributed array, say
the \( k \)th array on the line. Associated with this array is the
weighting function \( w_{jk}(t) \). The subscripts indicate that
\( w_{jk}(t) \) is the weighting function for the \( j \)th member of the set
of multiple ports associated with the \( k \)th array. An impulse
applied to a differential element of this array generates a
wave on the line with amplitude \( w_{jk}(\xi/c)u_{k}(t-\frac{z-\xi}{c})d\xi \),
where \( \xi \) is the coordinate of the element, \( u_{k}(t) \) is the impulse
response of the element, \( z \) is the coordinate at which the
amplitude is measured, and \( c \) is the propagation velocity of
waves on the line. The element generates waves propagating
in both directions along the delay line and the direction of
increase for \( \xi \) and \( z \) is chosen according to the direction of
propagation of whichever of these two waves is being con-
sidered. It is assumed that all elements of the array have the
same impulse response \( u_{k}(t) \). The amplitude of the wave at
the coordinate \( z \) due to the entire distribution of elements in
the launch array is

\[
a(t, z) = \int_{-\infty}^{\infty} w_{jk}(\xi/c)u_{k}(t - \frac{z - \xi}{c})d\xi.
\]

Next, suppose that \( z \) is the coordinate of a differential
member of the \( m \)th array. Associated with this array is the
weighting function \( w_{lm}(t) \). The subscripts indicate that \( w_{lm}(t) \)
is the weighting function for the \( l \)th member of the set of
multiple ports associated with the \( m \)th array. If an impulse
wave passes this element at time zero, the response of the
weighted element is \( w_{lm}(z/c)u_{m}(t)dz \), where \( u_{m}(t) \) is the impulse
response of the element. As before, it is assumed that all
elements of the \( m \)th array have the same impulse response
\( u_{m}(t) \), not necessarily equal to \( u_{k}(t) \). When the wave \( a(t, z) \)
passes this element, the response is

\[
w_{lm}(z/c) \int_{-\infty}^{\infty} a(\lambda, z)u_{m}(t - \lambda)d\lambda dz.
\]

The impulse response between the \( j \)th port of the \( k \)th array
and the \( l \)th port of the \( m \)th array is obtained by substituting
(12) in this expression and integrating the result over \( z \).

\[
h_{jklm}(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w_{jk}(\xi/c)w_{lm}(z/c) \\
\cdot u_{k}(\lambda - \frac{z - \xi}{c})u_{m}(t - \lambda)d\xi d\lambda dz.
\]

The right-hand side of (13) is the quadruple convolution

\[
h_{jklm}(t) = \left[ w_{jk} \ast w_{lm} \ast u_{k} \ast u_{m} \right](t),
\]

where \( w_{jk}(-t) = w_{jk}(t) \) is the time-reversed version of the
weighting function for the launch array and \( \ast \) denotes con-
volution. A factor \( c^2 \), which is a consequence of integrating
the time-domain weighting functions with respect to the
spatial variables \( \xi \) and \( z \), has been omitted from (14) since it
may be absorbed in the weighting functions. This expression
for the impulse response is valid for both continuous and
discrete weighting functions. For the discrete case, the con-
volutions with respect to the weights are sums instead of
integrals.

The frequency transfer function is given by the Fourier
transform of (14).

\[
H_{jklm}(f) = \mathcal{F}^{-1}\left[ W_{jk}(f)W_{lm}(f)U_{k}(f)U_{m}(f) \right],
\]

where \( W_{jk}(f) \) is the frequency response of the \( j \)th
member of the set of multiple ports associated with the \( k \)th
array.
where the upper-case symbols denote the Fourier transforms of the corresponding lower-case symbols and a bar over a symbol denotes complex conjugation. Thus, the response between the ports of a pair of arrays of a generalized transversal filter is equivalent to the response of four cascaded filters which correspond to the launch array reversed in space or conjugated in frequency, the receive array, a launch array element, and a receive array element.

$U_s$ and $U_m$ set the maximum possible bandwidth of signals to be filtered. If consideration is restricted to signals whose spectral components lie within the limits imposed by $U_s$ and $U_m$, the filtering effect of the array elements can be neglected. For discrete arrays, $U_s$ and $U_m$ also set the spacing required for contiguous elements. If the elements are so spaced and if consideration is restricted to signals with corresponding bandwidth or less, then the effect of the elements can again be neglected. Thus for both continuously tapped arrays and for continuous arrays, when the preceding conditions are satisfied, (14) and (15) reduce to

$$h_{jklm}(t) = [w_{jk}^{-1} \ast w_{1m}](t)$$  \hspace{1cm} (16)

and

$$H_{jkim}(f) = \overline{W_{jk}(f)} W_{1m}(f).$$  \hspace{1cm} (17)

From either (16) or (17), it can be seen that the impulse response $h_{jkim}(t)$ is the cross correlation between the receive array weighting function $w_{nm}(t)$ and the launch array weighting function $w_{jk}(t)$. As in the case of the Kallmann filter, if $u_k$ and $u_m$ serve as reconstruction filters, the discrete distributions of weights can be replaced by the continuous functions for which the weights are samples.

For a general signal $x(t)$, applied to the $j$th port of the $k$th array, the output at the $l$th port of the $m$th array is

$$y_{jklm}(t) = [x \ast w_{jk}^{-1} \ast w_{1m}](t).$$  \hspace{1cm} (18)

When the elements of both arrays are reciprocal, $y_{jk1m} = y_{1mjk}$, since in that case $h_{jkim} = h_{imjk}$.

The net effect of the additional complexity of the generalized filter is to cascade the filtering action of the launch array with that of the receive array. Thus, the analysis and synthesis of the generalized transversal filter are simple extensions of the analysis and synthesis for the Kallmann filter. In the synthesis procedure, the designer has a choice in the partitioning of a specified impulse response or frequency transfer function between the launch and receive arrays. Some implementations which demonstrate the advantage of this flexibility are described later.

**Propagation and Transduction**

The foregoing theory of transversal filters is based on the assumptions of nondispersive propagation and lightly coupled and nonreflecting transducer and tap elements. Since one of the most attractive features of this theory is the simplicity of transversal filter analysis and synthesis, these assumptions are desirable from a practical point of view. When they are not satisfied, the analysis and synthesis become difficult (e.g., [26]).

There are a number of other practical considerations in the implementation of transversal filters. The bandwidth and $TW$-product requirements of the particular application are significant factors in the choice of an implementation for a transversal filter. The bandwidth has an important relation to the physical size of the filter. For most applications, the length of an array in the filter is approximately equal to $cT/k$, where $k$ is the time-compression factor. Since $k = W_{max}/W$, where $W_{max}$ is the maximum possible bandwidth for the filter, the array length becomes $L = cTW/W_{max}$. Thus, for a fixed $TW$ product, the length of the array is inversely proportional to the filter’s bandwidth capability and large $W_{max}$ implies small size. Consequently, the filter’s mean frequency $f_0$ does not affect its length; it makes no difference whether a filter with a bandwidth of 100 MHz is centered at 100 MHz or at 1 GHz. Since nearly all materials have an acoustic attenuation that increases monotonically with frequency, and since the complexity of the peripheral electronics increases with frequency, small values of $f_0$ are desirable. Small values of $f_0$ to achieve low loss and simple electronics, with large values of $W_{max}$ to achieve small physical size, imply a large fractional bandwidth $W_{max}/f_0$. However, it becomes increasingly more difficult to obtain launch transduction with both low insertion loss and low reflectivity as the fractional bandwidth is increased. Thus, a compromise must be made. For most of the ultrasonic transversal filter (UTF) implementations to be described here, fractional bandwidths between 0.5 and 1 are feasible.

The ideal delay-line material for transversal filters has a dispersionless, low-loss mode of propagation and has the capability of wide-band low-loss transduction. The extent to which these various conflicting requirements can be met determines the success of a particular implementation. Some implementations which have proved to be successful are described below.

In the frequency region from 1 MHz to 10 MHz, satisfactory transduction, propagation, and tapping can be obtained with a wire delay line propagating torsional waves, using magnetostrictive transducers and taps; in the frequency region from 10 MHz to 500 MHz, they can be obtained with a single-crystal delay line propagating surface waves, usually employing piezoelectric transducers and taps; and in the frequency region from 500 MHz to 2000 MHz, they can be obtained with a single-crystal delay line propagating bulk waves, either longitudinal or shear, using thin film piezoelectric transducers and optoacoustic taps.

Each of these implementations is discussed in detail later. In addition to the combinations of delay-line materials, transduction, propagation, and tapping used for these implementations, there are many others that have been or could be used. Since propagation, transduction, and tapping have been studied extensively for delay-line applications, the reader is referred to the literature [27]-[30]. Some important results are summarized in this section.

There are two fundamental waves in unbounded isotropic solids: a dilational wave with propagation velocity $c_p = (\lambda + 2\mu)/\rho^{1/2}$, and a rotational wave with propagation velocity $c_t = \mu^{1/2}/\rho^{1/2}$, where $\lambda$ and $\mu$ are the Lamé constants of the material and $\rho$ is the density. The former is usually
called a longitudinal wave since the particle motion is in the direction of propagation; the stresses in the direction of particle motion are tensions and compressions. The latter is usually called a shear wave since the particle motion is transverse to the direction of propagation, thus causing shear stresses in the direction of particle motion. For both these waves, the propagation is nondispersive. The presence of boundaries complicates the situation, causing energy conversion between the two fundamental modes, with dispersion as the usual consequence. In some cases essentially nondispersive modes of propagation exist in bounded media. Of these modes, the most important for UTF implementations are 1) extensional waves in wires and ribbons whose dimensions, transverse to the propagation direction, are less than a wavelength, with propagation velocity \( c_e = \mu^{1/2}(3\lambda+2\mu)^{1/2}/\rho^{1/2}(\lambda+\mu)^{1/2} \), where \( E \) is Young's modulus; 2) torsional waves in wires whose diameter is less than a wavelength, with \( c_t = \mu^{1/2}/\rho^{1/2} \); 3) width-shear waves in sheets whose thickness is less than a wavelength but whose width is many wavelengths, with \( c_s = \mu^{1/2}/\rho^{1/2} \); 4) surface (Rayleigh) waves on plates whose thickness is several wavelengths, with \( c_R = \mu^{1/2}(0.87+1.12\sigma)/\rho^{1/2}(1+\sigma) \) where \( \sigma = \lambda/(2\lambda+\mu) \) is Poisson's ratio; and 5) longitudinal or shear bulk waves in rods whose transverse dimensions are many wavelengths, with velocities \( c_p \) and \( c_s \), respectively.

The attenuation in polycrystalline materials contains terms proportional to the first and to the fourth power of frequency. The first-power term represents the anelastic loss and for many materials decreases as the elastic modulus increases. The fourth-power term represents the Rayleigh scattering loss and increases with the size of the individual crystallites [30]. The scattering loss can be reduced in some wires and ribbons by the partial orientation of the crystallites resulting from cold working. In magnetostrictive materials, there are additional losses due to hysteresis and eddy currents [31], and at frequencies below 10 MHz these are dominant losses. Hysteresis losses depend on the magnetic state of the material and eddy current losses depend on the magnetostrictive coefficients and resistivity. Magnetostrictive delay lines operating in the torsional mode, at frequencies of about 1 MHz, can be constructed from wire with moderate resistivity such as NiSpan C, with losses as low as a few tenths of a decibel per meter, so that lines with as much as 30 meters of length and a TW product of \( 10^4 \) are feasible. An extremely simple and effective method of tapping such lines is described later.

A composite structure can be used to reduce the magnetic losses. A thin magnetostrictive film is formed, deposited, or plated on a nonmagnetic base wire. The base wire controls the propagation characteristics and the magnetostrictive film controls the tapping characteristics. This type of structure has been used with extensional waves [32], [33], and appears promising for applications in which the impulse response of the filter must be alterable. Another interesting composite structure consists of a thin magnetostrictive film vacuum deposited on an amorphous glass or fused silica plate [34]. The attenuation in these amorphous materials contains terms proportional only to the first and second powers of frequency, so that the propagation losses are acceptable at much higher frequencies than for polycrystalline materials. Wide-band delay lines operating in the width-shear mode at frequencies to 100 MHz can be constructed in this manner.

Amorphous dielectric plates can be used also for the propagation of surface waves [35]. This mode of propagation has received much attention in recent years, not only for amorphous plates but also for single-crystal dielectric plates [36]–[38]. Although the surface wave propagation losses in amorphous materials become excessive for frequencies above 100 MHz, single-crystal dielectric plates have been used at frequencies above 2 GHz. Surface waves are nondispersive only on the surface of a semi-infinite solid, but particle motion diminishes exponentially with distance away from the surface, and propagation on a plate only a few wavelengths thick is essentially nondispersive. Surface waves on plates more than a few wavelengths thick are the slowest of the nondispersive waves, having velocities between 0.87 and 0.96 of the shear wave velocity. Thus, for a fixed TW product and \( W_{\text{max}} \), an implementation of a UTF using surface waves as the mode of propagation will be smaller than an implementation using any other mode.

Surface waves can be guided in several ways [39], of which two are especially suitable for UTF implementations: a thin strip on the surface of a plate, with the propagation velocity of the strip lower than that of the plate; and a thin layer on a plate with a slot through the layer to the plate, with the propagation velocity of the layer greater than that of the plate. Both of these guides are dispersive. The slot guide is the more attractive of the two, because it has the least dispersion. There is a close analogy between guided surface waves and guided microwave electromagnetic waves on microstrips. Microstrip waveguide components such as directional couplers and hybrids have direct surface wave equivalents [40]. Surface waves have velocity and wavelength \( 10^3 \) times smaller than electromagnetic waves with the same frequency. Consequently, surface waveguide components are approximately \( 10^3 \) times smaller than their microstrip counterparts or conversely, for a fixed physical size, they can contain \( 10^3 \) times more components or be \( 10^4 \) times more complicated.

Amorphous and single-crystal dielectric materials are also used for the propagation of bulk waves, both longitudinal and shear, for frequencies from about 1 MHz to 100 MHz. As with surface waves, bulk wave propagation losses become excessive in amorphous materials for frequencies much greater than 100 MHz and single-crystal materials have been used for frequencies in excess of 2 GHz. At the higher frequencies, the wavelength is only a few microns and well-collimated acoustic beams with diameters on the order of 1 millimeter can be generated and propagated with essentially no dispersion, so long as the transverse dimensions of the medium are sufficiently larger than those of the beam.

In general, single-crystal materials are elastically anisotropic. Therefore, the orientation of the crystal axes with respect to the direction of propagation is an important consideration, since only in special directions will the direction
of energy flow be collinear with the normal to the wavefront [41].

Another aspect of propagation which is important in delay-line implementations is the propagation delay-time stability, especially with respect to temperature changes. The significance of delay-time changes for the tracking of a UTF matched filter was discussed earlier. Delay-time stability is also important in recirculating time compressors such as the DELTIC. Horologic steels have been developed which, in addition to being isopaustic, are magnetostrictive [42]. Wires made from these steels propagate torsional mode waves with good delay time stability and low propagation loss and can be tapped magnetostrictively. Although isopaustic glasses have been developed, they have large propagation losses relative to fused silica [43]; however, fused silica has a large temperature coefficient of delay and must be temperature controlled. Single-crystal quartz is an outstanding material in this context, even though its propagation losses at high frequencies are somewhat greater than for other crystalline materials in common use, since delay stable orientations of the crystal axes exist for shear waves both at room and elevated temperatures. Because of the importance of time delay stability to microwave acoustic devices, a considerable expenditure of effort could be justified in the search for additional crystalline materials with orientations that provide delay stability.

Although many types of interactions exist between acoustic waves and other forms of energy, the most commonly used types for the transduction of energy between electrical and acoustic forms are the electromechanical and magneto-mechanical interactions. Magnetostrictive, ferroelectric, and piezoelectric transducers are particularly important for UTF implementations. These transducers are reciprocal and can be used both to launch and to tap a wave. Another type—the optoacoustic interaction—is gaining importance as a means of tapping acoustic waves. In general, this type is not reciprocal and can be used only for tapping the acoustic beam.

Transduction in magnetostrictive materials is similar to that in ceramic ferroelectric materials. In both types, a bias field is required for linear operation, and a permanent field in the material can provide this bias. Some of the magnetostrictive materials and all of the ferroelectric ceramics commonly used for transduction can be self-biased.

There are two fundamental types of transduction in these materials: dilational, when the signal field is collinear with the bias field; and rotational, when the signal field is orthogonal to the bias field [44]. The collinear transducer is used for longitudinal and extensional waves and the crossed-field transducer is used for shear and torsional waves. Both magnetostrictive and ferroelectric materials are used extensively for these transducers. Both types of materials are used in collinear transducers for extensional waves in wires and ribbons and in crossed-field transducers for shear waves in strips and plates, and torsional waves in wires. When the medium itself is magnetostrictive, these various modes can be launched and tapped directly in the medium. Ceramic transducers can be bonded directly to the edges of ribbons, strips, and plates with some possible advantages relative to external magnetostrictive transducers in such applications. However, a good impedance match, which is sometimes difficult to obtain, must exist between the bonded transducer and the line; otherwise the transducer may have small fractional bandwidth [45].

Satisfactory theories for the operation of these transducers for all of the propagation modes involving linear particle displacement were developed many years ago. However, until a little more than 10 years ago, the various theories proposed for torsional transducers with circular cross sections were in serious disagreement with experimental observations [46]. In 1958, Yamamoto [47] provided an analytical description of such transducers which is in excellent agreement with experimental observations. Although the analysis was applied to magnetostrictive transducers, it can be extended to the ferroelectric case.

Single-crystal piezoelectric and ferroelectric transducers are used for all modes of propagation and at all frequencies, from subaudio through microwaves. At frequencies above 100 MHz the control of crystal thickness and satisfactory bonding techniques become critical problems [48]. The theory of these transducers is complicated because of their anisotropic electric and elastic properties. The search for new materials with improved properties is currently an active field.

Thin-film piezoelectric transducers, vacuum deposited on the propagation medium, are widely used for longitudinal and shear waves at frequencies above 500 MHz [49]-[51]. The most common materials are cadmium sulfide and zinc oxide, with the latter providing somewhat better insertion loss and bandwidth characteristics than the former. This advantage is offset to some extent by the fact that the vacuum deposition of zinc oxide is more difficult than that of cadmium sulfide. For both materials, the best performance is obtained by the application of thin-film impedance matching layers between the transducer and medium. As with the single crystal piezoelectric and ferroelectric materials, new materials with improved properties for thin-film transducers are being actively sought.

Optoacoustic interactions are used for the tapping of acoustic waves over a wide range of frequencies, from less than 1 MHz to more than 2 GHz. The tapping can be accomplished by the modulation of a light beam incident on a wave which is interior to or on the surface of an optically transparent acoustic medium. The light is spatially modulated by the acoustic wave, according to the phase and amplitude of the various spectral components in the acoustic wave and can be regarded as a continuously distributed tap array. Since neither the bulk wave nor the surface wave tap is generally reciprocal, they are not used for launch transduction. The theory of diffraction can be used to describe the behavior of the tap array and the resulting analysis usually depends on the detailed nature of the optical and acoustic arrangement.

Magneostricitive Filters

The torsional magnetostrictive wire delay line is unusually simple in concept and implementation. One or more turns of wire encircling the delay line can serve as either a launch
transducer or a tap with the weight controlled by the number of turns and their direction. The magnetic bias required for linear magnetostrictive transduction is obtained by passing a current down the delay-line wire, thus establishing a circumferential crossed-field bias within the delay line. If the delay-line material is sufficiently remanent, the current can be removed and a permanent bias field will remain. When a current is passed through a launch wire, a longitudinal magnetic field is established in the delay line which acts as the signal field. The bias field and signal field interact with the magnetostrictive properties of the delay line to create a torsional strain which propagates along the line (the Wiedemann effect [52]). Conversely, a torsional strain in the delay line interacts with the bias field and the magnetostrictricion to create an additional longitudinal field component (the inverse Wiedemann effect) and as the strain propagates along the line and passes a tap, this associated field induces an EMF in the tap. Yamamoto gives the relationship between the torsion per unit length $\theta$, the longitudinal magnetostrictive coefficient $\lambda_t$, and the magnetic field intensities $H_L$, $H_o$, and $H_k$ as

$$\theta = 3\lambda_t(H_o)H_L^2H_o/rH_k^2,$$  \hspace{1cm} (19)

where $H_k^2 = H_1^2 + H_o^2$, $H_1$ and $H_o$ are the longitudinal and circumferential field intensities, respectively, and $r$ is the radius of the delay-line wire [47]. Equation (19) indicates that when $H_1 \ll H_o$, the torsion is linear in $H_o$.

Torsional delay-line UTFs have been developed and fabricated by Whitehouse and Lindsay, operating at center frequencies in the region of 1 MHz, with almost 100 percent fractional bandwidth and with $TW$ products of approximately $10^5$ [25]. The $TW$ product of these filters can be increased to approximately $10^4$ by compensating for propagation losses, and it may be possible to increase the operating frequency to about 10 MHz without decreasing the fractional bandwidth, by employing a delay-line wire of smaller diameter and redesigning the tap structure. However, advances beyond these values are blocked by the fundamental limitations due to the polycrystalline nature of the delay-line wire.

The torsional delay-line UTF is especially suitable for the implementation of multiple-port arrays. In one implementation, a single turn tap is passed through a small linear ferrite core and serves not only as a tap, but also as the primary winding of a wide-band transformer. The weight of the tap is determined by the direction and number of the secondary turns on the core and the summation is accomplished by connecting the secondary windings of all the taps in series. If the core at each tap has a large enough aperture to permit a number of secondary windings, then a corresponding number of independent weighting functions and multiple ports can be obtained. In actual practice, such a filter might be constructed by passing a shuttle, carrying a continuous wire, through each core in turn, in a manner analogous to the stitching operation in sewing. The tap weights are determined by the direction and number of times the shuttle is stitched through each core before passing on to the next one. Each time the shuttle is stitched through the complete set of cores, it carries a wire, the two ends of which constitute a port. Thus, the number of independent ports possible is limited by the number of wires that can be passed through the core apertures without blocking the passage of the shuttle. A single-port filter of this type is shown schematically in Fig. 9 and a section of an eight-port filter operating at a center frequency of 1 MHz, with large fractional bandwidth capability, and with a maximum $TW$ product of 512 is shown in Fig. 10 [25].

The number of possible independent ports can be increased by using split cores. In this case, the secondary winding are prefabricated from a continuous wire and then slipped over one section of each of the split cores, after which the other core sections are placed in position to close the cores. Such a multiple-port filter is shown schematically in Fig. 11. The fabrication process for the secondary windings can be performed on a loom. The warp would consist of $2n$ wires, where $n$ is the maximum weight for any tap in the filter; half of the wires for positive weights and half for negative weights. At each position along the wires corresponding to the tap positions, $|w_i|$ wires are lifted — $|w_i|$ from the positive group for a positive weight, or $|w_i|$ from the negative group for a negative weight, where $w_i$ is the weight of the $i$th tap. At each tap position, then, a number of wires corresponding to the magnitude of the weight at that tap position has been lifted from the group of wires corresponding to the sign of the weight. The warp is then gathered into a compact bundle of wires, care being exercised to keep the appropriate wires lifted. The split core halves are then inserted into the spaces between the lifted and undisturbed wires and the other halves are used to close the cores. Finally all of the wires are connected in series, with the negative wires bucking the positive wires. A tapping structure results, with the correct weights. For multiple ports, $2nm$ wires are used, where $m$ is the number of ports, with the wires connected in groups of $2n$ for each port. A Jacquard loom, such as the one used at the Massachusetts Institute of Technology for the fabrication of "braid" memories, can be used to loom these tapping structures [53].

The torsional delay line UTF described above has fixed weighting functions. Another type of wire delay-line UTF, using extensional waves, has an alterable weighting function. This filter makes use of a nondestructive-read, ferroelectric memory [54] which uses the coincidence between a high intensity extensional wave pulse propagating along the delay line and a time varying current in the delay-line wire to form a remanent field in the delay-line material. The magnitude and sign of the remanent field at a particular point on the line are controlled by the magnitude and sign of the current in the delay-line wire at the time the acoustic pulse passes the point. An acoustic pulse launched along the line induces an EMF in the delay-line wire given by $h(t)=w(t)$, where $w(c/c)$ is the function describing the remanent field. Thus, the voltage appearing across the ends of the delay-line wire is the weighted and summed output from contiguous taps, with the weighting function $w(t)$. To replace the weighting function with a new one, a time varying current of the desired form is passed through the wire, and an acoustic pulse is launched along the delay line. A UTF of this type can be made to operate at a frequency of 1 MHz, and 100 percent bandwidth with a $TW$ product of $10^5$ might be possible. The
The $TW$ product in this filter is limited by the fact that the high intensity acoustic pulse needed to set the remanent field drives the material of the delay line into a nonlinear acoustic region, with large propagation losses as a consequence.

Another ferroacoustic delay line, originally developed as a stress-current coincidence memory [33], has interesting possibilities as a delay line for UTFs when used in a current-current coincidence mode similar to that used for plated wire memories. The delay line consists of a beryllium-copper wire, plated with a thin permalloy film. The remanent field is set by the coincidence of a current in the delay-line wire and a current in a wire perpendicular to the delay line. The impulse response is related to the remanence by the same equation as for the stress-current coincidence filter: $h(t) = \mu(t)$. Since in this filter the remanence is not set acoustically, the limitations on the $TW$ product caused by the rapid attenuation of the acoustic setting pulse are absent and it should be possible to construct filters operating at 10 MHz, with 100 percent bandwidth and with $TW$ products of approximately $10^4$.

A UTF with multiple-output ports as well as alterable weighting functions is possible if the wire delay line is replaced by a glass or fused silica plate with a vacuum-deposited thin magnetostrictive film on its surface. As with the wire delay lines, the state of remanence can be set by either stress-current or current-current coincidence. If the acoustic transducer is many wavelengths across, the acoustic wave will be well collimated and plane wave propagation will occur. Thus, the acoustic stress will be constant in each differential strip transverse to the direction of propagation. If a number of wires parallel to the direction of propagation are placed in juxtaposition with the surface, each of these wires can be used as an independent channel for setting the remanence in that part of the film directly under the wire. Once the remanence has been set, each wire becomes an independent output port. The number of ports possible is limited only by the spacing required between channels to prevent crosstalk, and by the width of the plate. A similar filter, in which the magnetostrictive film has uniaxial anisotropy, limits the weighting function to binary values [34]. Filters operating at frequencies as high as 100 MHz, with 100 percent fractional bandwidth and with a $TW$ product of $10^4$ might be possible.

**Surface Wave Filters**

One of the most interesting advances in microwave acoustics is the use of surface waves on piezoelectric and ferroelectric substrates [55]. Although surface waves can be used at frequencies from below 1 MHz to above 1 GHz, the most promising applications for UTFs are at frequencies from
about 10 MHz to 500 MHz. For these frequencies, surface wave devices are small in size and are easily fabricated by techniques similar to those used by the semiconductor industry for large scale integrated circuits.

Many of the piezoelectric materials used as substrates are also semiconductors, and amplification of surface waves on such materials has been achieved by the interaction of the electric field of the surface wave with charge carriers drifting at the surface wave velocity [56]. Since these fields extend above the surface, a nonsemiconducting substrate can be used in conjunction with a piezoelectric semiconductor: the two materials need only to be brought into proximity with a drift field applied to the semiconductor [57]. This separation of the functions of acoustic propagation and semiconductor amplification allows each material to be optimized.

Surface waves on isotropic nonpiezoelectric substrates have been studied extensively [35]. Although transduction on such substrates is difficult, there is still considerable interest in them due to their superior quality for guided surface waves [58]. The most successful transducer for nonpiezoelectric substrates is the Sokolinskii comb transducer [59]. This transducer consists of a longitudinal mode piezoelectric crystal, bonded to one face of a buffer plate whose opposite face has periodic grooves and is in contact with the substrate. Electrical excitation of the crystal produces a periodic stress distribution on the surface in contact with the comb, due to the grooves in the plate. This technique results in an efficient surface wave transducer with good directivity, and with the direction of surface wave propagation normal to the grooves.

For piezoelectric materials the White-Voltmer interdigitated electrode transducer has been highly successful. This transducer consists of an array of periodically spaced parallel conductive strips on the surface of the substrate, with alternate strips electrically connected together to form a comb-like array of interdigitated electrodes [60]. When the array is excited, periodic stresses are generated at the surface by the interaction of the resulting electric field with the piezoelectric material, and a surface wave is launched. The Sokolinskii transducer must have a large number of grooves for low insertion loss and has an impulse response which is approximately a finite duration sine wave, with the number of cycles equal to the number of grooves. Consequently, it is a narrow-band transducer. The White-Voltmer transducer requires a large number of strips for low insertion loss, when the piezoelectric coupling coefficient is small, and has an impulse response similar to that for the Sokolinskii transducer. Thus it is also a narrow-band transducer. However, when the coupling coefficient is large, low insertion loss can be achieved with a few electrodes, and in this case the transduction is wide-band [61].

The surface wave transducer array is an example of a distributed array of the type described in the discussion of generalized transversal filters. If the input array has weighting function $w_i$ and the output array has weighting function $w_o$, then the impulse response of the array pair is $w_i * w_o$, which is the cross correlation of the input and the output array weighting functions. Transducer arrays of the types described above, when used as input and output arrays, have an impulse response (the cross correlation of two sine waves of equal duration) which is a sinusoid with a triangular envelope with a duration twice that of the impulse response of the individual arrays. Thus, an array pair acts as a narrow-band filter whose bandwidth, for a fixed array length, is inversely proportional to the number of strips or grooves and whose frequency transfer function is of the form $\sin(\pi x/x_0^2)$.

The result indicates that a spectrum analyzer can be constructed with high resolution and low sidelobe response. If more than one pair of such electrode arrays are cascaded, then the total frequency transfer function will have the form $\sin(\pi x/x_0^r)$, where $r$ is the number of pairs in cascade. For $r=1$ the first minor lobe is down 26 dB, and for $r=2$ the first minor lobe is down more than 50 dB. In addition, when two sets of such cascaded transducer pairs are deposited on the same crystal substrate and driven from separate sources, the phase difference between the two signals can be measured simultaneously with their spectra [62] since differential phase stability is assured by the use of a common substrate.

In many applications it is desirable to have frequency transfer functions which are either broad-band or have small sidelobes. Two types of weighting functions which might be used to achieve these responses are period variation weighting, where the spacing between interdigitations changes along the array, and equally spaced interdigitations with amplitude modulation of the individual tap elements. An essentially rectangular wide-band frequency response can be achieved with the former, if the impulse response of the array has an instantaneous frequency which changes linearly with time; a sidelobe-free frequency response can be achieved with the latter, if the amplitude modulation is Gaussian.

A third type of weighting function, in which the interdigitated transducer elements are phase modulated, is required for many applications. An example of a phase-modulated code with a narrow correlation function and with uniformly small sidelobes is a Barker code [63]. A transducer array using simple electrode pair inversion to implement a 7-bit Barker code is shown in Fig. 12(a), along with the corresponding electric fields. Electrode pair inversion, however, fails to provide the impulse response expected of a Barker code, because of the equipotential regions formed at the phase changes of the code, with consequent loss of spatial components in the stress distribution.

A phase-coded transducer which has the required stress distribution has been developed by Whitehouse and Lindsay, and utilizes one additional electrode with a reconnection of the other electrodes. This transducer and its electric fields are shown for comparison in Fig. 12(b). By appropriately weighting the amplitude of a transducer composed of these phase-invertible elements, an arbitrary impulse response can be synthesized by the techniques given in Section I.

As an example of the synthesis of an arbitrary impulse response, a pulse-in to pulse-out distributed surface-wave transducer delay line is described. It is possible to obtain pulse-in to pulse-out behavior if the input and output arrays
are each a simple electrode pair. However, a large number of elements in each array is required to achieve good signal-to-noise ratios on substrates with low electromechanical coupling coefficients. Arrays with large numbers of elements can be designed for pulse-in to pulse-out operation, but care must be taken to choose array weightings whose correlation functions have low sidelobes in order to avoid signal interference when continuous signals are transmitted.

The Huffmann codes [64] are phase-reversal codes with impulse-equivalent behavior. They exist for all code lengths, and have only one small correlation sidelobe on each side of the central impulse. However, these codes are not attractive for surface wave delay lines because they require a large variation of the array element amplitudes, with a consequent difficulty of fabrication. This problem can be alleviated by the use of a pair of binary phase-reversal codes which have the property that the sum of the correlation functions of the code pairs is an impulse, with no side lobes. Golay has studied such codes, which he calls “complimentary series.” For their properties, the reader is referred to his work [65]. Since these codes are phase-reversal codes, the three-electrode differential transducer is a natural choice. In phase-coded arrays that have these complimentary series as weighting functions, the number of positive electrodes is not equal to the number of negative electrodes. Consequently, the transducer is capacitively unbalanced. However, the correlation function of the negative of a function is the same as that of the function itself. Thus, one of the pair of codes can be negated, so that a parallel connection of the two transducers will have capacitive balance. When the weighting functions are so chosen, the transducer can be conveniently driven by a wide-band differential transformer, with some gain in rejection of interference. A delay line using a 128-element Golay complimentary series array connected for capacitive balance is shown in Fig. 13. The implementation of complimentary series phase-coded surface-wave transducers should make possible wide-band pulse delay lines, with negligible sidelobe interference, for operation to 500 MHz. These delay lines should have TW products of about $10^4$ and they can be fabricated with existing photoresist technology. These delay lines, when combined either with microwave acoustic amplification or microelectronic amplification, can be used for transversal filter time compressors or for DELTICs, as well as for matched filters and spectrum analyzers.

**Optoacoustic Filters**

With the recent advent of lasers, interest in the optoacoustic interaction as a means of tapping the information contained in an acoustic delay line has increased. Typically, the information in the acoustic beam will be contained in a band of frequencies about a central carrier frequency $f_0$. From spatial sampling considerations, the acoustic beam can be considered to be composed of contiguous segments of length $\Delta_m/2 = c/2f_m$, where $c$ is the acoustic propagation velocity, $f_m$ is the largest frequency by which the carrier is modulated, and $f_m \leq f_0$. Each of these segments will interact with that portion of the incident light beam which it intercepts. If the incident light beam intercepts a length $\Delta_m/2$ of the acoustic beam, then there is only one tap on the delay line, and the output of that tap will simply be modulated in time as the acoustic waves propagate by. If the incident light beam intercepts a length larger than $\Delta_m/2$, then each of several contiguous segments of the acoustic beam will interact separately with that portion of the incident light beam which it intercepts, so that several contiguous optoacoustic taps result, each of which can be separately weighted in amplitude, phase, or polarization.

If the acoustic beam is a bulk wave propagating in a liquid or a solid, the delay experienced by any given light ray traveling through the acoustic beam will depend on the degree to which the density along its path has been increased or decreased by the alternating compressions and rarefactions of the acoustic beam. Not only will the delay along the optical path be alternately increased and decreased by such action, but also the ray will be refracted or polarized by the acoustic medium, resulting in both phase and amplitude corrugations of the optical wavefront. Such corrugations will generate spatial interference patterns similar to those caused by a conventional diffraction grating.

---

Footnote:

7 The application of complimentary series to delay lines was proposed by J. M. Speiser, and has been investigated in detail by Speiser and Whitehouse; a paper describing this work is in preparation.
A grating-like optical interference pattern can also be generated using acoustic surface waves. Such waves can change optical path delay in a manner analogous to the bulk wave interaction just described [66], and they can also be used to frustrate the total internal reflection of a light beam at a solid–air interface. For the latter effect, a second solid is placed adjacent to the first, and the surface wave is propagated on either of the two adjacent surfaces so that a maximum frustration occurs at the crests of the wave, and a minimum frustration occurs at its troughs. This gives rise to a reflected and a transmitted light beam both of which are modulated in amplitude across their widths.

In many cases of interest only the light diffracted into the first-order lobe(s) is important, and the light in other orders, including the zeroth order, can be neglected or can be eliminated by means of spatial filters or polarization analyzers. In general, the rays diffracted into the first-order lobe will travel at some angle relative to the acoustic wavefronts and thus will have the possibility of overlapping with rays being diffracted by an optical tap contiguous to the one in which the ray originated. The proportion of overlapping rays decreases as the width $d$ of the acoustic beam is decreased, or as the maximum modulation frequency $f_m$ is decreased. For acoustic bulk waves, decreasing the acoustic beamwidth will usually cause a decrease in the amount of light diffracted, and decreasing $f_m$ reduces the bandwidth, so that a tradeoff between acoustic beamwidth, bandwidth, and signal-to-noise ratio may be required. Large fractional bandwidths on the order of 50 to 100 percent are not, in fact, possible when $d \gg \Delta \omega/\lambda$, where $\Delta \omega = c/j_0$, and $\lambda$ is the optical wavelength (this is the region where Bragg diffraction dominates Raman-Nath diffraction [67]). The interaction of light with surface waves is inherently broad-band, since for surface waves, the transverse dimension $d$ is of the order of $\lambda_0$.

An optical system can be regarded as a two-dimensional analog of an electrical system, with the input and output planes of the optical system corresponding to the input and output ports of the electrical system. It is known that the optical counterparts of various electrical filtering operations (e.g., low-pass, high-pass, bandpass, bandstop, differentiation, integration) can be performed by means of lenses and apertures [68]. The information to be filtered is presented at the input plane as a transmission or reflection function with an arbitrary distribution of amplitude, phase, and polarization, and the input plane is illuminated with coherent monochromatic light. Spatial multiplication occurs when a second transmission or reflection function is placed either in the same plane as that occupied by the first function, or at an image of that plane. An optical multiplier based on these principles is illustrated in Fig. 14.

When the transmission function moves across the input plane, then the resulting distribution in the output plane will correspondingly change in time in a manner depending upon the intervening optical system. If the intervening optical system serves to spatially multiply the input transmission function by a second transmission function (the weighting function), then the resulting spatial product is equivalent to a weighted tapping of the input transmission function. Optical integration of this weighted tapping forms an optical transversal filter.

If the optical integration is astigmatic so that it integrates this product only along the direction of motion, then a multiplicity of simultaneous transversal filters can be formed. In particular, if the input plane is divided into a number of strips parallel to the direction of motion, then each strip and its corresponding output can be independent of all other strips and outputs.

An example of an optical transversal filter is a film strip moving along an axis in the input plane. Unfortunately, film-processing delays and slow film movement make this form of optical filter too slow for high-speed real-time signal processing. High-speed and real-time capability, however, are available in acoustic delay lines in which light interacts photoelastically with the acoustic waves in an optically transparent medium. Although acoustic delay lines do not have the high two-dimensional information storage capability of film, a large number of them can be used side-by-side, filling up the whole input plane, to partially offset this loss.

Fig. 15 shows two examples of such an optoacoustic filter in which the electrical signal of interest has been transformed into an acoustic signal to be interrogated by the light beam from the left. The modulated light beam is then weighted by (or compared to) a reference function as represented by a transparency or phase plate or coded polarization sheet [Fig. 15(a)], or by another acoustic signal in a second delay line [Fig. 15(b)], and finally integrated and presented at the output of a light detector. This second representation illustrates the ability of the optoacoustic filter to program the
reference function in real time. The positions of the signal and reference planes can be interchanged, but their separation (or the separation between one plane and the image of the other) must be small enough to prevent diffraction overlap of adjacent optical information cells.

Fig. 16 is the diagram of a device utilizing two acoustic cells oriented to take advantage of Bragg-angle incidence of the optical beams. One cell is placed in juxtaposition to the other cell (or its image) to rediffract the first-order lobe transmitted through the polarizer $P_1$. When the two acoustic signals are matched, the net phase change of any light ray undergoing the double diffraction is constant, regardless of the acoustic frequency, so that the wave passing through $P_1$ is plane. When the signals are not matched, the optical wavefront passing through $P_2$ may be a plane wave of lesser intensity, a plane wave in a different direction, or a non-plane wave, so that the intensity at the focal point of the lens is below the maximum given by matched signals. A slightly Doppler-shifted version of the reference signal will produce an output plane wave which will focus to a different point in the focal plane of the lens. If the polarizers $P_1$ and $P_2$ are deleted, some other means (e.g., spatial filtering, or optical heterodyning) must be used to separate the twice diffracted beam from the zero-order (undiffracted) light collinear to it. Squire and Alsup [18] have constructed a device of this type which can operate at a frequency of 1 GHz with 50 percent bandwidth and a $TW$ product of $10^7$.

A modification of the two-cell Bragg filter described above is presently under investigation and provides a simple means of separating the undiffracted light from the signal-carrying diffracted light. The two configurations being investigated are shown in Fig. 17. In the first, the two cells are placed so that the axes of acoustic propagation are oriented at the angle $\theta - \phi$, where $\theta$ is the Bragg angle for a carrier frequency $f_1$, and $\phi$ is the Bragg angle for a carrier frequency $f_2$. In either configuration, the beam of light which is diffracted by an unmodulated carrier $f_1$ in the first cell will be incident at the appropriate angle to be rediffracted by an unmodulated carrier $f_2$ in the second cell in the direction specified by the Bragg conditions for the two cells. Furthermore, an unmodulated carrier frequency $f_1 + \Delta f$ in the first cell will diffract a Bragg beam which will be rediffracted by an unmodulated carrier frequency $f_2 + \Delta f$ in the second cell into the same direction indicated above for frequencies $f_1$ and $f_2$; the plus sign is appropriate to Fig. 17(a), and the minus to Fig. 17(b). In both cases, the twice-diffracted light deviates from the direction of the incident light by approximately twice the Bragg angle.

If one of the acoustic cells in the preceding example is replaced by a fixed phase or amplitude grating designed to represent the signal to be matched, then a device results such as that shown in Fig. 18 [69], where the loss of versatility provided by an alterable weighting function has been offset by a gain in light signal intensity.

A single acoustic cell can act as a Fresnel lens while maintaining a unique Bragg diffraction geometry when the signal to be processed is a linear FM sweep, or "chirp" [70], [71]. As diagrammed in Fig. 19, this configuration satisfies the Bragg diffraction condition simultaneously for all frequencies in the acoustic beam at only one instant of time, and it
Fig. 19. Optoacoustic transversal filter matched to a linear FM signal ("chirp").

Fig. 20. Optoacoustic transversal filter utilizing frustrated total internal reflection of light by surface waves

is at that instant that the acoustic beam focuses the diffracted portion of the diverging input light to a single point without the use of auxiliary lenses. An alternative form of this technique uses a collimated incident light beam and anisotropic polarization rotation of the focused diffracted beam [72]. It should be possible to achieve bandwidths and $TW$ products in this type of filter comparable to those given in the two previous examples.

The frustrated total internal reflection by means of surface waves, described earlier, is utilized in the filter shown in Fig. 20 [73]. The beam diffracted from the first surface wave is spatially filtered within an afocal lens system which images the first surface wave upon the second surface wave. The twice-modulated light is gathered by a lens to a point at which is placed a photodetector, whose output represents the correlation, or convolution of the two acoustic signals propagating as surface waves. Like the first Bragg device described, this filter has the capability of an alterable weighting function, but its bandwidth limitations are governed by realizable carrier frequencies, instead of beamwidth. Such a filter should have performance characteristics comparable to those of the surface-wave filters described earlier.

CONCLUSIONS

Linear transversal filters are ideal for the implementation of filters for processing complicated signals with large time-bandwidth products, because of the simplicity of the synthesis procedure for the transversal filter. Although the implementations described in this paper as well as other existing UTFs provide heretofore unavailable signal processing capability, there are important immediate applications for which even greater capacity is required. Microwave acoustic technology, which has provided the basis for many implementations given here, also shows great promise for the implementation of filters to meet present and future advanced requirements. Because of the close relationship between the theories and techniques employed in the fields of acoustic microwaves and electromagnetic microwaves, increased collaboration among the workers in these two fields will make significant contributions in the development of these advanced devices.

REFERENCES
