Microsonics

J. L. Abita and H. K. Charles, Jr.

The field of microsonics and in particular surface acoustic waves (SAW) has developed rapidly during the past ten years. From the initial theoretical work on geological tremors by Lord Rayleigh (1885) and from the decades of work during this century on bulk acoustic wave devices for oscillators, filters, and delay lines, SAW-microsonics has emerged as a viable working technology that can delay, shape, limit, mix, and convolve signals over a frequency range from 1 MHz to several GHz. This article presents the theoretical basis, design development, and applications theory for SAW devices. It traces the history of SAW activity at the Applied Physics Laboratory and identifies areas of technological potential.

I. Introduction

Microsonics is a science concerned with the microscopic deformation of an elastic solid by propagating wavelike disturbances. These disturbances are classified as "bulk waves" if the motion is confined within the bulk of the medium or as "surface waves" if the motion is localized on the boundary or "skin" of the medium. Microwave acoustics suggests a mechanism similar to that resulting in sound propagation.

Anyone who has observed ripples on a pond or geological tremors is familiar with surface wave phenomena. Similar effects may be created on the surface of solids by local disturbances that cause waves to spread throughout the bulk and along the surface. However, important physical differences exist between the macroscopic analogies mentioned and what are considered microscopic surface acoustic waves (SAW) on solids.

In contrast to the relatively large displacements of the medium usually observed in the case of water and seismic waves, the wave motions in the case of SAW disturbances are microscopic because the displacements are on an atomic or molecular scale measured relative to an equilibrium "lattice" (a fixed reference system and not necessarily a crystalline structure). The total SAW disturbance is an integral effect of a large number of small (atomistic) individual displacements.

The use of SAW at ultrasonic frequencies (10 MHz to 1 GHz) on a special class of materials called piezoelectrics is giving rise to new families of devices for storing and recognizing signals, separating one signal from another, generating signals, and performing complex mathematical signal processing functions.

SAW devices are of interest because (a) the sound wave travels 10^5 times slower than a corresponding electromagnetic wave, (b) the energy in the wave is available for tapping everywhere

along the device surface, (c) the devices are compatible with the planar surface processing methods associated with microcircuit technology, (d) the power density is high, and (e) the acoustic losses are low compared to electromagnetic transmission lines of similar delays. Typical SAW velocities are on the order of 3×10^5 cm/s. Thus a SAW delay of 3 μs would require 1 cm of crystal as compared to a kilometer of air transmission line for the same electromagnetic delay.

An electrical signal is converted to an acoustic wave by means of an interdigital transducer (IDT). Its operation is based on the expansion and contraction of a piezoelectric material (e.g., lithium niobate or quartz) in response to an electric field. The electric field is created by applying a potential difference to two parallel interleaved (i.e., interdigital) thin-film metal electrodes (fingers) deposited on the crystal surface. A time-varying electrical input causes the crystal to vibrate and generate a sound wave along the surface. The sound wave may be converted to an electrical signal (i.e., inverse piezoelectric effect) at the other end of the crystal using a second IDT. The finger width and the spacing between fingers are related to the acoustic wavelength at the frequency of operation.

2. Theory

Analytical Basis

In an attempt to describe geological tremors, Rayleigh[1] developed an analytical formulation for waves traversing the free surface of an infinite, homogeneous, isotropic, elastic half space. In 1967 Viktorov[2] presented a generalized theory of Rayleigh and Lamb waves that is considered the foundation work for microsonics.

In the simple model adopted by Rayleigh, the particle displacements normal, W, and parallel, U, to the free surface (described above) are given in steady-state form by

$$U = Ak\left[e^{-qz} - \frac{2qs}{k^2 + s^2}\, e^{-sz}\right] \sin(kx - \omega t)$$

$$W = Aq\left[e^{-qz} - \frac{2k^2}{k^2 + s^2}\, e^{-sz}\right] \cos(kx - \omega t),$$

where k is obtained from the characteristic equation,

$$(k_t/k)^6 + 8(k_t/k)^4 + 8[3 - 2(k_t/k_t)^2](k_t/k)^2 - 16[1 - (k_l/k_t)^2] = 0,$$

z and x are the normal and longitudinal coordinates (the surface is taken to be infinite in extent along the x and y axes with the propagation vector in the positive x direction), and k_t and k_l are the transverse and longitudinal wave numbers, respectively, $q^2 = k^2 - k_l^2$ and $s^2 = k^2 - k_t^2$.

Solutions of the characteristic equation or Rayleigh equation give the wave velocity, V, in terms of the measurable material parameters. It has six roots, only one of which corresponds to a real material, i.e.,

$$(k_t/k) = V/V_t = \frac{0.87 + 1.12\sigma}{1 + \sigma},$$

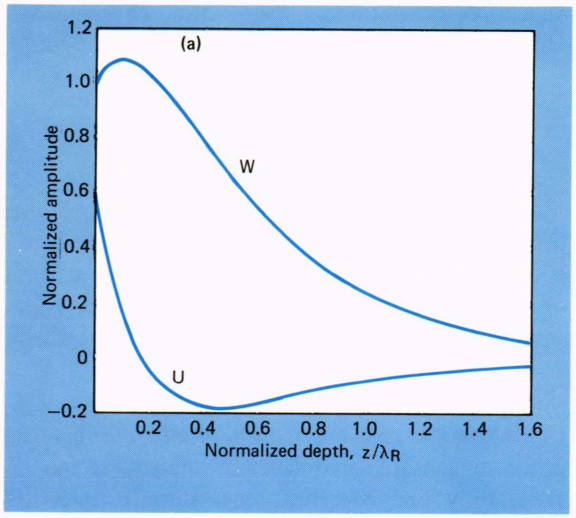
where σ is an elastic constant characterizing the propagation medium (Poisson's ratio) with the limits of 0 and 0.5. As σ varies between these limits, the Rayleigh wave-phase velocity increases monotonically from 0.87 V_t to 0.96 V_t where V_t is the velocity for a bulk transverse wave. Of particular importance is the nondispersive nature of the waves; i.e., the velocity and phase are not functions of the frequency.

In Fig. 1a the normalized amplitudes of U and W are plotted as a function of normalized depths from the surface. From this plot, details of particle motion as shown in Fig. 1b may be ascertained. At the surface, particles rotate on an elliptical trajectory with the direction of rotation being retrograde to the propagation direction. At increasing depth, the magnitude of both displacements is diminished; the longitudinal component (U) is 0 at 0.18 λ_R ($\lambda_R \equiv$ Rayleigh wavelength $= 2\pi/k$). At this point the wave is purely transverse; below this depth the longitudinal component magnitude increases, returning the particles to an elliptical path. However, because of a sign reversal in U, the direction of rotation is reversed. Finally, at approximately two Rayleigh wavelengths into the medium, both components of motion are essentially zero; thus the wave (energy) is confined to a surface layer about $2\lambda_R$ in thickness.

This analysis provides a reasonable description of surface waves for isotropic media that allow elastic lattice-point deformation. Media having

[1] Rayleigh, *Proc. London Math. Soc.* **17**, 1885, 4.

[2] I. A. Viktorov, *Rayleigh and Lamb Waves*, Plenum, New York, 1967.



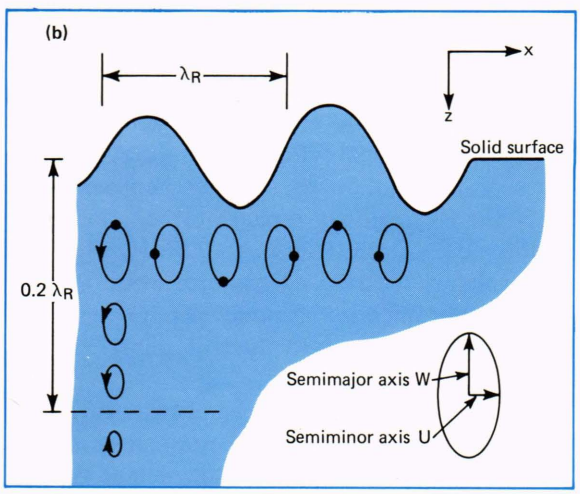


Fig. 1—Rayleigh wave particle displacement.
(a) Amplitudes normal (*W*) and parallel (*U*) to the free solid surface of propagation as a function of depth into the solid.
(b) Trajectories as a function of depth into the solid.

other than ideal properties are of practical interest. In general, such materials are nonhomogeneous, anisotropic, and elastic to a proportional limit, and they display vibrational/electromagnetic interactions. A reasonably accurate analytical description is impossible in the most general case. We must restrict ourselves to practical materials that, to a high degree, are homogeneous and isotropic with or without a restricted class of interactions (i.e., nonpiezoelectric or piezoelectric, oriented single-crystal conductors, semiconductors, or nonconductors). This encompasses most materials now being investigated.

A fundamental annoyance associated with surface waves is the concomitant generation of bulk waves.[3] Bulk waves, such as sonar, are a familiar concept. A solid can support similar "internal" elastic modes and oscillation of the entire medium (plate modes). In SAW devices, if the interference caused by the generation of bulk waves is ignored, unexpected device behavior may result; thus an accurate model should incorporate bulk modes. Bulk wave generation, interference, and elimination have been investigated, but the usual approach is to neglect their presence unless spurious responses and interference make elimination mandatory.

Modeling

A practical method for surface wave generation or transduction relies on coupling to a crystalline lattice via electromagnetic fields in piezoelectric materials. In its elemental form, an IDT as shown in Fig. 2 is used. The IDT is a comb-like structure composed of a thin-film metallic overlay delineated as interwoven fingers. An applied potential difference between combs stresses the surface lattice. A time-varying potential difference will cause time-varying stresses that travel from the transducer in accordance with the spectral content of the applied potential difference. Since this technique is used in the majority of SAW components being developed and employed, we will limit our discussion to the modeling of IDT's on piezoelectric materials. Table 1 lists the properties of the primary piezoelectric materials being used in the manufacture of the device.

In modeling for a particular application, two complementary approaches are useful. The first answers the question, "Given a desired signal processing function, which structure(s) will approximate this behavior?" The second answers the question, "Given the structure, what physical and electrical characteristics can be expected?" The latter question relates to device implementation with input-output terminal properties specified.

Those having some familiarity with SAW technology, perhaps by exposure to review articles,[4] may have the impression that each device has physical and electrical uniqueness. However, upon close examination one finds that most of them can be described by a common model for electrical parameterization and response. If we consider

[3] J. deKlerk and R. M. Daniel, *Appl. Phys. Lett.* **16,** 1970, 219.
[4] M. G. Holland and L. T. Clairborne, *Proc. IEEE* **62,** 1974, 582.

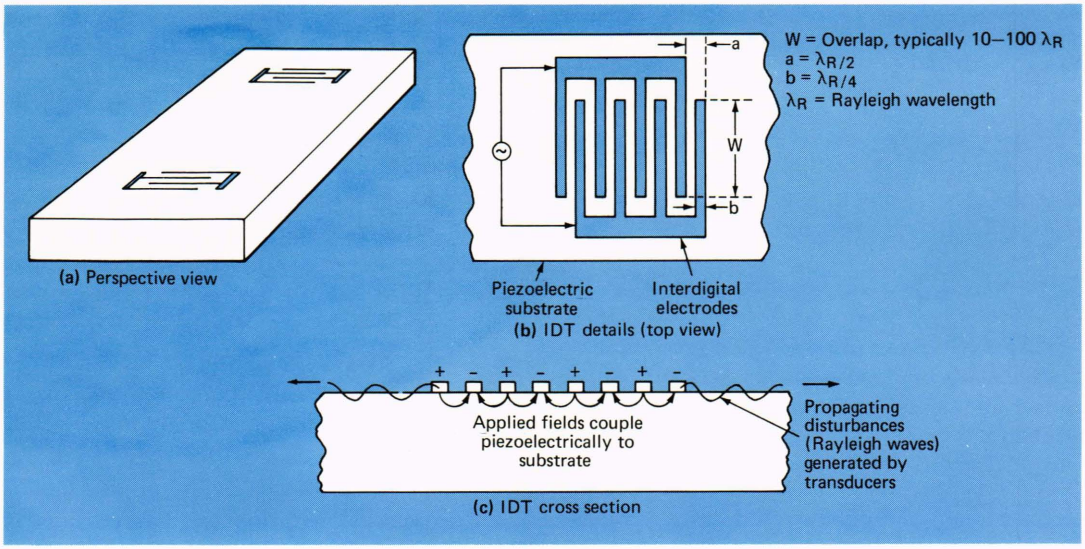


Fig. 2—Basic interdigital transducer (IDT) delay line.

only metallic overlays on piezoelectrics, they fall under the analytical structure of transversal filters.[5] In a transverse filter, a signal is passed along a delaying medium with taps distributed along its length. The output from each of the taps is weighted in amplitude and/or phase and then is summed to give the output.

The frequency response of an apodized IDT (an IDT with varied finger overlap) may be related, at least to first order, to the Fourier transform (FT) of a continuous weighting function. To determine the expected behavior of a transducer, one need only calculate the FT of the apodization or weighting function. Conversely, given the frequency response, a transducer can be designed via the inverse FT.

Electrical Characterization

A first-order equivalent circuit, Fig. 3, represents one section of an interdigital structure. Originally used by Mason[6,7] to describe bulk oscillations, it offers a reasonable (experimentally verified) model for surface wave devices. Voltages and currents correspond to lattice point forces and velocities, respectively. Second-order effects (next-nearest-neighbor electrode couplings, regeneration,

Table 1

MATERIALS (PIEZOELECTRIC)

Material	*Crystal Cut*	k^2 (%)	V_r (m/s)	τ (μs/cm)	*TCD* (ppm/°C)	*Loss* (dB/μs)
$LiNbO_3$	YZ	4.3	3485	2.87	85	1.6
$LiTaO_3$	ZY	0.74	3329	3.12	37	—
$Bi_{12}GeO_{20}$(BGO)	(111,011)	1.7	1680	5.95	125	1.5
SiO_2	YX	0.23	3159	3.15	−22	4.1
SiO_2	ST–X	0.16	3150	3.17	<3	—
ZnO/Al_2O_3	XZ	0.28	2740	3.79	40	—
ZnO	BASAL	1.0	2700	3.70	40	—
AlN/Al_2O_3	XZ	0.63	6170	1.48	—	—
PZT	BASAL	≈4.3	2200	4.55	—	≈5 (PZT 6 @ 40 MHz)
BeO	YZ	—	6580	1.52	—	—
CdS	XZ	0.62	1718	5.82	—	—
Tourmaline	XZ	0.28	4400	2.27	45	1

k^2 = piezoelectric coupling coefficient V_r = Rayleigh wave velocity τ = delay per unit length TCD = temperature coefficient of delay

[5] C. Atzeni and L. Masotti, *IEEE Trans. Microwave Theory Tech.* **MTT-21,** 1973, 505.

[6] W. P. Mason, *Electromechanical Transducers and Wave Filters,* 2nd ed., Van Nostrand, Princeton, NJ, 1948.

[7] D. A. Berlincourt, D. R. Curran, and H. Jaffe, *Physical Acoustics,* W. P. Mason (ed.), Academic Press, NY, 1964.

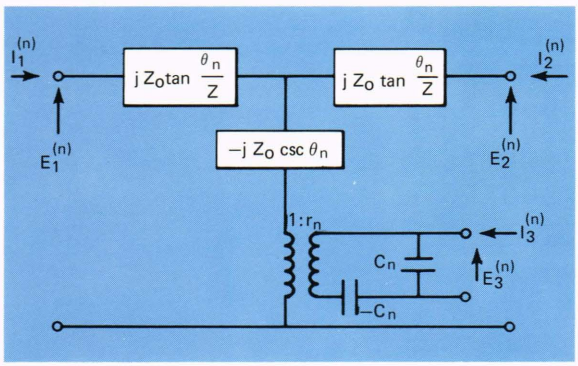


Fig. 3—First-order equivalent circuit of one section (nth) of an interdigital structure. Three-port model: two acoustic (ports 1,2) and one electric (port 3).

bulk wave leakage, etc.) have been neglected. These effects can be incorporated into a more complex model requiring machine computation.

The equivalent circuit chosen is one dimensional, whereas surface fields are two dimensional. This ambiguity is resolved by further simplification of the physical problem into either a crossed-field or in-line-field representation. The differences between the two representations have been given elsewhere[8] and, briefly, relate to the dominant electric field component responsible for Rayleigh wave transduction. The negative capacitance, $-C_n$, will be short circuited for the crossed-field approximation. The choice of crossed-field or in-line model is usually determined experimentally by examining the nature of the particular piezoelectric used. For relatively weak electroacoustic coupling and a small number of transducer periods, differences between the two models become negligible so that either one gives a reasonable description of port properties. However, there are materials and orientation for which neither model gives meaningful results.

The admittance matrix defines a relationship between currents and voltages in the form

$$I_j^{(n)} = \sum_{k=1}^{3} Y_{jk}^{(n)} E_k^{(n)}.$$

For the nth section, the admittance matrix in the crossed-field approximation is given by

$$Y_{ik}^{(n)} = \begin{bmatrix} -j \cot \theta_n & j \csc \theta_n & -jr_n \tan \theta_n/2 \\ -j \csc \theta_n & j \cot \theta_n & -jr_n \tan \theta_n/2 \\ -jr_n \tan \theta_n/2 & -kr_n\theta \tan_n/2 & j(\omega C_n + 2r_n^2 \tan \theta_n/2) \end{bmatrix},$$

where $\theta_n \equiv \frac{2\pi f}{V}$, r_n is the transformer ratio given by

$$r_n = (-1)^n \sqrt{\frac{V}{L_n} C_n k^2 Z_0},$$

L_n is the electrical transmission line length, C_n is the capacitance of the one electrode section, k^2 is the effective piezoelectric coupling constant for surface waves, and Z_0 is a characteristic impedance related to the material constants of the medium. Equal conductive line and gap width is assumed in deriving this transformer ratio.

The single-section admittance matrix above represents a first-order approximation neglecting several secondary effects. The matrix, despite its drawbacks, is the basic starting point for a proper electrical characterization of any given interdigital structure operating on a piezoelectric medium. The overall response is obtained by joining individual sections with their acoustic ports in series and their electrical ports in parallel, and, with the aid of a high-speed computer, calculating the overall admittance matrix.

This procedure not only provides the electrical port parameters of a transducer but also yields the frequency response. The use of Fourier transformations, while providing the expected frequency response, does not electrically characterize the device or consider source and load impedance implicit in the admittance matrix approach.

3. Design Considerations

Transducer Patterns

Assuming a two-transducer filter, Fig. 2, one must consider the overall frequency response as the combined individual responses of the two separate transducers. In the simplest arrangement, one transducer is apodized and the other is a broadband array consisting of a few finger pairs (Fig. 4a). In this configuration, only one transducer dictates the overall frequency response making design implementation easy, but the filter suffers from high insertion loss and undesirable rounding caused by the $\sin(2\pi f)/2\pi f$ frequency response of the uniform transducer. These disadvantages may be skirted by using a pair of apodized transducers in conjunction with a multistrip coupler[5] (Fig. 4b). The multistrip coupler allows each transducer to operate essentially independently with the overall response simply the product of the individual responses. Multistrip couplers are only practical on

[8] W. R. Smith et al., *IEEE Trans. Microwave Theory Tech.* **MTT-17**, 1969, 856.

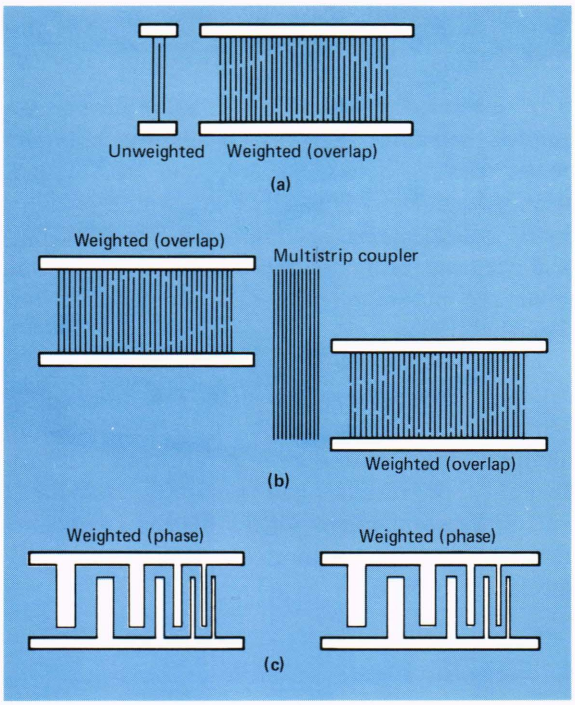


Fig. 4—Possible two-transducer filter-weighting configurations.

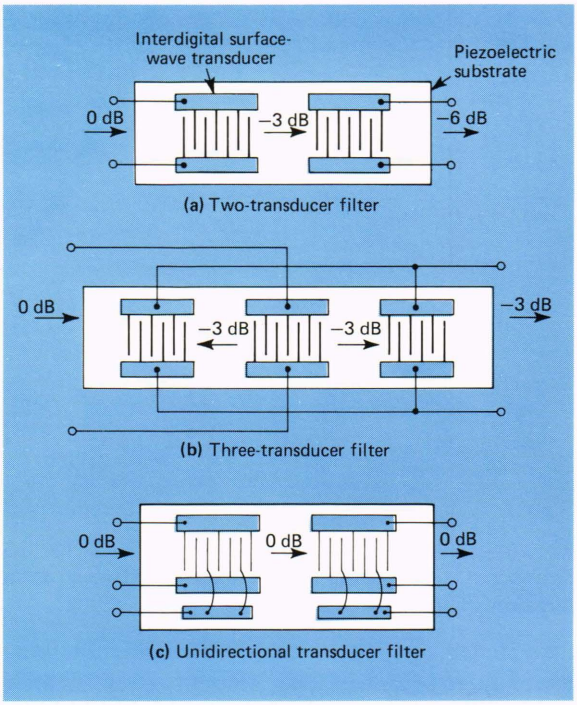


Fig. 5—Multiple transducer filters.

high coupling substrates. Figure 4c illustrates a phase-weighting method of achieving the desired frequency response. Phase modulation (variable finger spacing including finger removal or withdrawal) is used with uniform finger overlap, thus allowing both transducers to be weighted without the use of a multistrip coupler.

Two-transducer filters, shown in Figs. 2 and 5a, suffer an inherent insertion loss of 6 dB (in addition to any other loss mechanism) due to the bidirectional nature of the two transducers. The input and output transducers are symmetric about their centers; thus the input transducer radiates only half the power toward the output transducer, resulting in a 3-dB loss. Reciprocally, the output transducer can reconvert only half the acoustic energy incident on it from one side into electrical output, resulting in another 3-dB loss. Half of this loss can be removed by placing a second output transducer on the substrate on the opposite side of the input transducer as shown in Fig. 5b. Bidirectional loss can be completely eliminated by using a multiphase unidirectional transducer (Fig. 5c). The multiphase drive removes the bidirectional symmetry and permits complete conversion from electrical signals to acoustic signals traveling in one direction.

Materials

Currently the primary SAW materials being used are quartz (SiO_2) and lithium niobate ($LiNbO_3$). These are relatively inexpensive, easily cut and polished, and consistent in their properties. Lithium niobate is used for its high coupling coefficient, while quartz's primary attribute is its temperature stability.

The curves in Fig. 6 give the minimum insertion loss as a function of fractional bandwidth ($\Delta f/f$) for several piezoelectric materials. As long as the electrical Q, given by $(\pi/4k^2)\Delta f/f$, where k^2 is the piezoelectric coupling coefficient, is less than the reciprocal of the fractional bandwidth, one can easily match the device across the entire bandwidth, and the minimum insertion loss is limited to transducer bidirectionality. This condition is expressed as:

$$\frac{\Delta f}{f} \leq \sqrt{\frac{4k^2}{\pi}}.$$

For lithium niobate, $k^2 = 4.3$, which restricts the fractional bandwidth to 25% for matching considerations. If the fractional bandwidth of the transducer is greater than allowed by the above criterion, the matching circuit must be loaded or

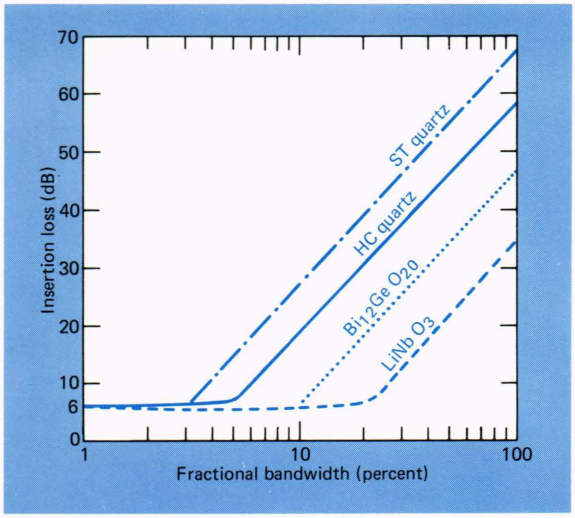


Fig. 6—Minimum achievable insertion loss for bidirectional transducers on various substrates as a function of fractional bandwidth.

mismatched so that its bandwidth will not limit device performance. The amount of mismatch required for larger fractional bandwidths increases at a rate of approximately 12 dB per octave of bandwidth. The insertion loss curves shown in Fig. 6 are optimum values for the two-transducer filter configuration (Fig. 5a) below center frequencies of 500 MHz (above 500 MHz other loss mechanisms enter). For a three-transducer configuration (Fig. 5b), the curves would be reduced uniformly by 3 dB, and for unidirectional transducer configurations still another 3-dB (Fig. 5c) improvement can be had.

In Fig. 6, $LiNbO_3$ appears to be an excellent material for most devices; however, other effects limit its usefulness. The temperature coefficient of delay is 85 ppm/°C compared to quartz with a range of 0 to ≈25 ppm/°C. Because of its high coupling coefficient, $LiNbO_3$ is more subject to distortion effects such as reflections and regenerations. Other parameters that may affect substrate choice include propagation velocity, beam steering, and beam spreading. Low-velocity substrates minimize size for a given time delay, while high-velocity materials result in reduced resolution requirements during the fabrication process. In devices with large time delays (long distances between input and output transducers), beam steering and spreading effects become very important. Quartz generally tends to spread the acoustic beam faster than predicted by diffraction theory, but $LiNbO_3$ for a particular crystal orientation (YZ-cut) is almost self-collimating.

Matching

In general, low-insertion-loss SAW devices require a matching network to assure optimum power transfer between source and load impedances. The networks, usually made with lumped or distributed elements, accomplish the required impedance transformation and in some cases provide input and output filtering of unwanted spurious signals.

The use of matching networks partially negates the main advantages associated with SAW phenomena, namely, size reduction and temperature stability. However, the ease of fabrication, reproducibility of results, and reliability associated with SAW components (including their matching networks) outweigh minor compromises for many special applications.

4. Fabrication

The SAW medium must be carefully prepared before the components are made. Each crystal must be optically flat (at least on the propagation surface) and have accurately aligned and defined axes relative to the crystal boundaries. The substrates are then prepared for the transducer pattern delineation.

Metallized patterns are defined by using standard microelectronic processes. In the basic, most widely used technique, metal is deposited over the entire surface of the polished surface. The metallized surface is coated with photoresist (a light-sensitive polymer) by spraying or spinning. The photoresist is exposed to UV light through a mask containing the required interdigital pattern. The exposed resist areas are removed during developing, while the remaining photoresist protects the desired interdigital pattern against the metal etch. After chemically etching, the protective photoresist is removed by stripping in a solvent. Using this technique, devices up to 400 MHz (2.0-μm line widths) have been successfully fabricated.

Another technique, referred to as reverse lift off, uses photoresist as a mask. The substrate is first coated with photoresist. The pattern is defined by illumination through a mask, and the holes for the metal are developed in the photoresist coat. The metal is deposited through the holes in the photoresist. The excess metal (on top of the photoresist) and the photoresist itself are removed

by stripping. No metal etches are used, and lines as fine as 1 μm may routinely be defined using this technique. To achieve the sharp line definition required, the mask must make intimate contact with the photoresist. Flexible masks and special fixturing are usually required to make this technique work.

Optical photomasks are limited to dimensions of about 1 μm because the wavelength of radiation necessary to cause the exposure of the photoresist (typically 0.4 μm) produces diffraction patterns comparable to device dimensions. One-micrometer resolution would limit standard SAW operation to approximately 900 MHz. Electron beam exposure can be used to overcome these difficulties of diffraction and resolution. A modified form of reverse lift off is used in which a special resist coated on the substrate (e.g., polymethyl methacrylate) is exposed by a computer-controlled flying spot electron gun. In effect, the holes for the metal are "drawn" into the resist. After exposure and hole developing, the metal (aluminum) is deposited, and the excess along with the resist is stripped away.

Dynamic focusing and computer-controlled substrate stages have done much to reduce cost and to process large areas by step and repeat operations. Using electron beam masking, line widths less than 100 nm are routinely achieved. SAW reflective array structures require fine grooves cut into the crystal surface. These are formed by ion-beam bombardment or etching.

Ion-beam etching was developed as a tool for surface finishing and fine polishing in which material is very slowly removed (typically 1 nm/min) by momentum transfer from a beam of accelerated argon ions. The beam strikes the surface at a relatively small angle with respect to the surface, removing all microcracks and protuberances in its path. All materials may be polished by ion-beam etching, but the rate of etch depends on the beam energy and the particular material under ion bombardment.

To form the groove patterns for SAW reflective arrays,[9] it is necessary to mask the surface of the device substrate. One method is to coat the surface with a thin layer of evaporated aluminum, approximately 500-nm thick. Openings in the aluminum for the grooves are formed by standard photolithography, and the substrate is then exposed to the ion beam. The beam removes material from the substrate surface through the mask as well as part of the mask. Groove depth is determined by time of exposure to the ion beam. This process is highly controlled, since beam energy is easily monitored and the exposure times are long enough (several minutes) to allow accurate time measurements.

Nonuniform groove depth is obtained by exposing different portions of the device to the ion beam for different times. Usually, the ion beam is relatively collimated (typical beamwidth is 1 to 2 mm), and the substrate is translated past the beam on a controlled stage driven by a programmable stepping motor. Once the program is developed and the ion beam voltage and current are well controlled, accurate reproducibility is achieved.

5. General SAW Applications

SAW may offer a convenient way to realize components. However, designing to a specific response function, although straightforward, is usually complex, requiring many iterations to finalize the design. Once the design is established, its repeatability is ensured by the reliability of microelectronic fabrication processes and techniques. Current SAW devices being implemented in various systems are delay lines, bandpass filters, pulse compression (expansion) filters, coded filters, "matched filters," and oscillators. Convolvers, programmable filters, amplifiers, phase shifters, acousto-optical modulators, and chirp transform processors are still under development.

Delay Lines

All SAW devices are basically delay lines. They usually consist of two IDT arrays separated by a distance, l, on a piezoelectric substrate (Fig. 7a). The time delay, t_d, for the device is given by

$$t_d = l/V,$$

where V is the Rayleigh velocity of the acoustic wave for the given substrate material (Table 1). It is apparent that the maximum time delay for the V of a given material will be determined by the spacing between IDT's and hence substrate size. Substrate size is dictated by the constraints of the circuit package size and the availability of large piezoelectric crystals. $LiNbO_3$ crystals 15 cm long have been prepared and, assuming that 10 cm of the length could be used for delay (the rest for

[9] R. C. Williamson and H. I. Smith, *Electron. Lett.* **8,** 1972, 401.

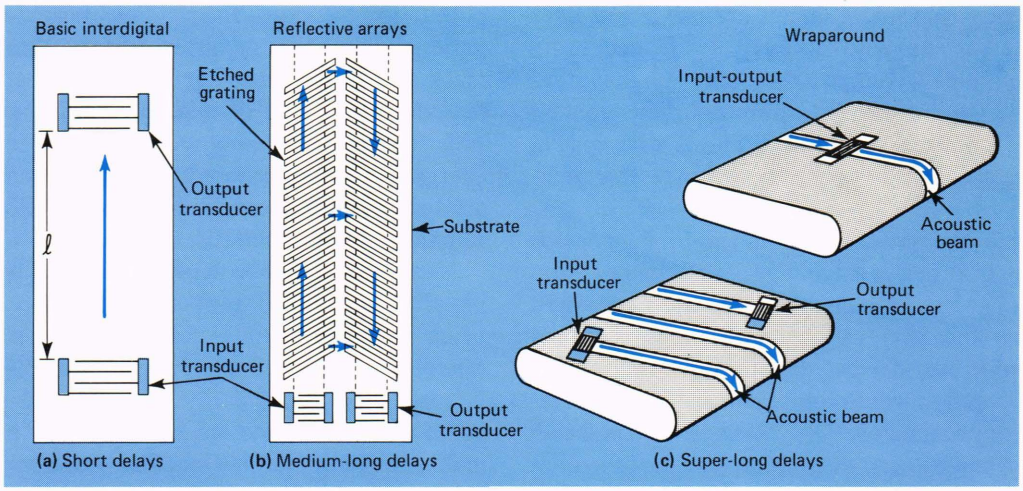


Fig. 7—SAW delay line structures.

crystal polishing, IDT's, and absorbing materials), the resulting delay would be 28.7 μs.

To achieve greater delay without increasing physical size and cost (large single crystals are expensive), other techniques have been developed. The most successful is the reflective array[9] (Fig. 7b), which operates as follows. An acoustic wave is generated by a wideband IDT and propagates toward one of the reflective arrays of grooves that have been ion-beam etched in the substrate surface. Each groove reflects acoustical energy at right angles to the incident beam; the frequency and amplitude of the reflected wave depend on the periodic groove spacing and the number and depth of the grooves, respectively. A second array parallel to the first reflects the acoustic wave toward an output IDT.

On $LiNbO_3$, reflective array delays in the 50- to 60-μs region are readily achieved using available crystal lengths. On quartz, which is another popular piezoelectric that is available in long single crystals and has greater delay coefficient per centimeter, delays of 100 μs have been obtained.[5]

For even larger delays, experimental wraparound delay lines (Fig. 7c) have been constructed.[10] The closed-loop path (Fig. 7c top) uses one transducer for both input and output, and the surface wave makes multiple passes. Long delays suffer from aliasing and loading problems produced by reflected signals generated by the IDT during each pass of the surface wave. With angled input and output transducers (Fig. 7c bottom), aliasing and loading effects are minimized, and the helical path followed by the surface wave produces long time delays with little distortion. Using this structure, delays in the 1- to 10-ms range have been observed with minimum loss in the signal energy, provided that the corners and end faces are optically polished and lapped.

IDT Dispersive Delay. Surface waves exhibit no natural velocity dispersion; however, dispersion may be "created" by allowing the distance of propagation to vary with frequency.[11] Dispersive delay lines are fabricated the same way as are uniform delay devices except that the spacing between electrodes of each IDT is not held constant. Instead, a "graded periodicity" (Fig. 8) makes each electrode pair resonant at a different fixed frequency and, since the physical position of each electrode is different, the time delay (i.e., distance between synchronous electrode pairs divided by the acoustic velocity) will vary with frequency. The input IDT is then dispersive and, if the output transducer is delineated so that the highest frequencies travel the shortest distance (Fig. 8a), the dispersion can be doubled. When an impulse (short compared to a half period at the highest frequency) is applied to the input (left side) of the device shown in Fig. 8a, an expanded frequency-modulated pulse is detected (after a time delay equal to the distance traveled by the highest frequency divided by the acoustic velocity) at the output (right side of Fig. 8a). This expanded pulse is called a "downchirp" because the delay versus frequency slope is negative. The

[10] G. S. Kino and H. Matthews, *IEEE Spectrum* **8,** 1971, 22.

[11] J. deKlerk, *Phys. Today* **25,** 1972, 32.

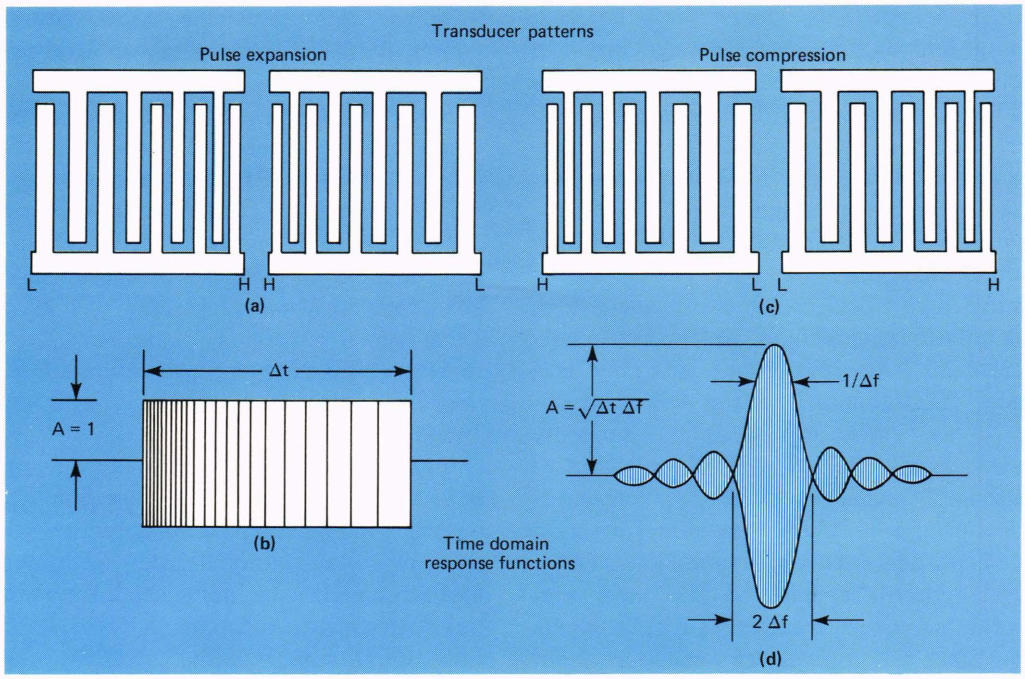


Fig. 8—SAW interdigital dispersive delay lines.

length of the expanded pulse (Fig. 8b) is

$$\Delta t = (d_L - d_H)/V,$$

where d_L is the propagation distance for the lowest frequency (f_L) and d_H is the propagation distance for the highest frequency (f_H). Δt is called the dispersion. The bandwidth, Δf, of the expanded pulse is

$$\Delta f = (f_H - f_L).$$

If the expanded pulse is fed into the input transducer of Fig. 8c, a compressed pulse having a $\sin(2\pi f)/2\pi f$ response is received at the output transducer (Fig. 8c). This pulse compression is referred to as an "upchirp" since the delay versus frequency slope is positive. The amount of compression is determined by the pulse compression ratio (PCR), which is the product of the bandwidth and pulse length (time-bandwidth product),

$$\text{PCR} = \Delta f\, \Delta t.$$

The amplitude A of the compressed pulse (Fig. 8d) is

$$\text{A} = \sqrt{\text{PCR}} = \sqrt{\Delta f\, \Delta t},$$

and the 3-dB bandwidth of the compressed pulse is $1/\Delta f$.

Reflective Array Dispersive Delay. The reflective array compressor (RAC) consists of an input and output IDT (broadband, few fingers, whose spacing is determined by the center frequency of the device) and two sets of obliquely placed etched gratings (the grating angle depends on the ratio of SAW velocity in the initial propagation direction to the velocity in the reflected direction (Fig. 9)).[9] The spacing between slots in

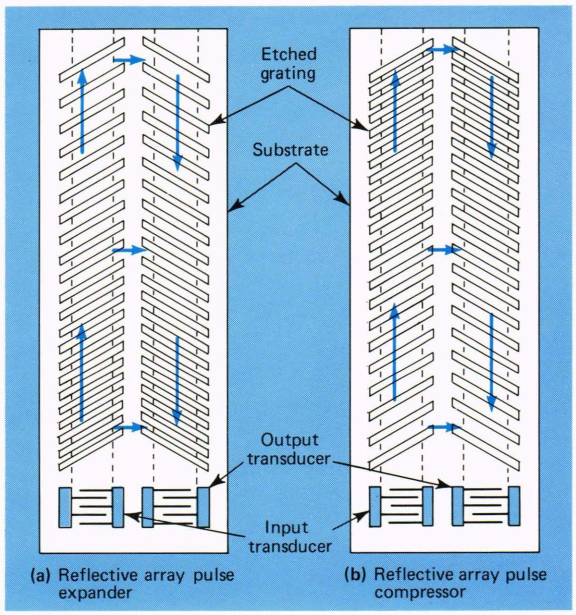


Fig. 9—SAW reflective array dispersive delay lines.

the gratings is varied, the closest being adjacent to the input and output grids. The structure will selectively reflect frequencies according to the spacings between slots. Thus, high frequencies will be reflected nearest the transducers and low frequencies farthest away, yielding short propagation paths for high frequencies and long paths for low with the response amplitude synthesized by groove-depth weighting.

Reflective array devices have extremely high tolerance to fabrication defects. Defects cause only minor problems, since the groove acts only as a mechanical discontinuity and the reflection coefficient of the entire groove is not seriously affected by a small defect. The IDT's on reflective arrays are simple and easily fabricated, the groove-to-groove spacing in the propagation direction being one acoustical wavelength compared to one-half wavelength for a standard IDT and one-quarter of an acoustical wavelength if split fingers are used.[5] Since the grooves are located at about 45° to the propagation directions, the required photographic resolution is 0.707 of that required for the standard IDT.

Reflective arrays have two disadvantages: a larger insertion loss (because acoustical reflection at each groove must be kept small to avoid dispersion) and one extra fabrication step, i.e., ion-beam etching.

Table 2 presents the latest state-of-the-art figures for IDT and reflective array dispersive delay lines.

Bandpass Filters. SAW bandpass filters are used in military communications and radar systems for both IF and RF applications.[5] To date the only large-volume commercial application has been filters for television tuners.[9] Table 3 presents a representative list of specifications available from SAW bandpass filter technology today along with a projection for the next 2 to 5 years. Figure 10 illustrates a representative filter and its transducer pattern in addition to its transducer apodization pattern.

Despite problems associated with matching, bulk wave generation, and insertion loss, SAW bandpass filters offer some attractive advantages over most conventional filter technologies (10 MHz to 1 GHz), including small size, low weight, low-cost construction, simplicity, high reproducibility, and design flexibility.

A filter's size essentially depends on its bandwidth; for a bandwidth greater than a few megahertz it occupies an area of less than 100 mm^2 and a thickness of a few hundred μm. Even when mounted, the weight of such a device is only several grams. Given a polished and oriented crystal, the electrode pattern is the only structure that must be accurately defined. The pattern is

Table 2

TYPICAL PERFORMANCE CHARACTERISTICS OF SAW DISPERSIVE DELAY STRUCTURES. CW INSERTION LOSSES FOR BOTH REFLECTIVE ARRAYS AND INTERDIGITAL STRUCTURES ARE HIGH, RANGING FROM 25 TO 50 dB

Center Frequency (MHz)	*Bandwidth, Δf (MHz)*	*Dispersion, Δt (μs)*	*Time-Bandwidth Product ($\Delta f\ \Delta t$)*
Interdigital Structures			
30	5	2	10
30	5.6	50	280
60	20	50	1 000
100	35	5	175
150	50	2.5	125
300	100	10	1 000
1250	500	0.5	250
1300	500	0.5	250
Reflective Array Structures			
60	6	20	120
60	2.5	125	312
60	6	100	600
200	50	30	1 500
500	250	40	10 000
1000	512	10	5 120
1000	200	50	10 000

Table 3

SAW BANDPASS FILTER PERFORMANCE PARAMETERS, CURRENT AND NEAR-TERM PROJECTIONS (2–5 YEARS)

Parameter	*Current Typical*	*Current State of the Art*	*Projected*
Center Frequency, f_0	10 MHz–1.0 GHz	5 MHz–1.5 GHz	1 MHz–2 GHz
Bandwidth	50 kHz–0.4 f_0	50 kHz–0.4 f_0	20 kHz–0.8 f_0
Minimum Insertion Loss	6 dB	0.7 dB	0.5 dB
Minimum Shape Factor	1.2	1.2	1.2
Minimum Transition Bandwidth	50 kHz	50 kHz	20 kHz
Sidelobe Rejection	45 dB	65 dB	70 dB
Ultimate Rejection	60 dB	80 dB	80 dB
Deviation from Linear Phase	±1.5°	±1.5°	±1.0°
Amplitude Ripple	0.5 dB	0.5 dB	0.05 dB
Triple-Transit Suppression	−40 dB	−50 dB	−50 dB

controlled by highly reproducible photolithographic procedures after the initial photomask has been produced. Any desired number of identical units can be produced, and the filters require no tuning after manufacture.

Within limits, wide flexibility in design is possible. The phase and amplitude of filters may be controlled separately; i.e., a wide range of amplitude responses can have linear phase. The independence of amplitude and phase is a result of the signal delay between the electrodes in the transducers. Filters that incorporate delays in this manner are called transversal filters.[5] SAW provides a very convenient way to implement small high-frequency transversal filters.

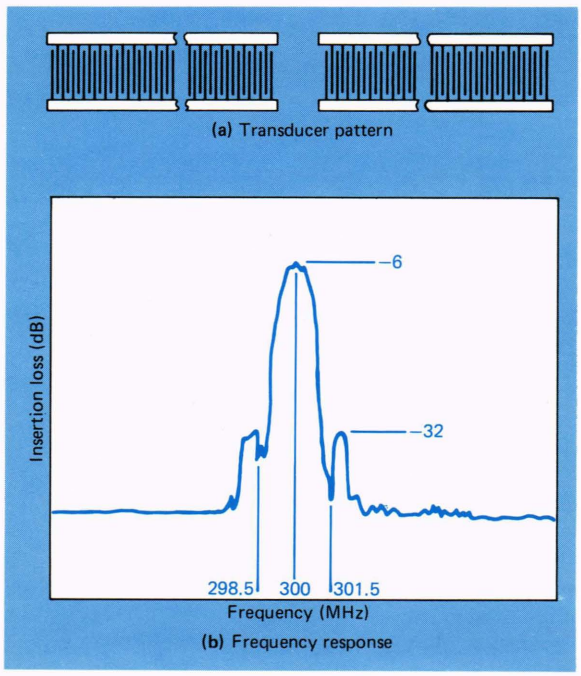


Fig. 10—Response of a SAW bandpass filter.

Resonant Structures

Narrowband filters ($\Delta f/f < 1\%$) can be implemented by using an IDT with many fingers, since the acoustic bandwidth is inversely proportional to the number of fingers.[12] Narrow bandwidths with quality factors of 10^4 and 10^5 would require long transducers resulting in poor yields. Other approaches for obtaining narrowband filters have been considered, such as long transducers with finger withdrawal, wraparound delay, and resonant "cavities."

Long Transducers with Finger Removal. This approach uses long transducers such as those that would have a large number of fingers for a given narrowband response (e.g., Fig. 10) except that large groups of fingers are omitted from each transducer. An example of the finger-withdrawal technique is shown in Fig. 11a. A two-transducer filter design with 2404 fingers per IDT has all except 32 fingers per ID comb removed. The bandwidth of the original filter (with no fingers removed) at a center frequency of 100 MHz is approximately 0.08% (given essentially by $1/\tau_t$, where τ_t is the total time for a SAW to travel the IDT). As fingers are removed, more peaks appear in the frequency response (Fig. 11b). The frequency interval between the peaks is $1/\tau_p$, where τ_p is the SAW propagation time from elemental comb to elemental comb. By the proper selection of elemental group spacings, it is possible to obtain a frequency response with one or at most a few peaks.

[12] R. H. Trancrell and M. G. Holland, *Proc. IEEE* **59,** 1971, 393.

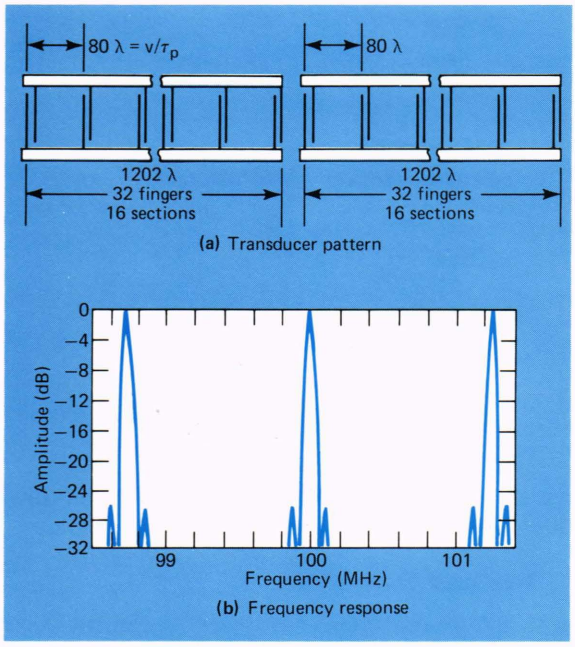


Fig. 11—SAW comb filter generation by finger removal.

Wraparound Delay. Wraparound delay lines (Fig. 7c) may be used as a narrowband comb filter. Here the frequency envelope is determined by the elementary comb response, while the frequency separation is given by $1/\tau_d$, where τ_d is the SAW propagation time around the path. Peak shape (width) is a function of loss and is given approximately by $1/\tau_e$, where τ_e is the time after which the amplitude has decreased to $1/e$ of its original value. Combinations of transducers can be used to isolate response peaks and provide a narrowband filter.[10]

Resonant Cavity. SAW crystal resonators or resonant cavities are a relatively new class of SAW devices. The basic resonator structure is shown in Fig. 12a. It consists of two SAW grating reflectors with an IDT placed inside the cavity to provide electrical input/output coupling. The operating characteristics of the cavity will depend on the reflection coefficient of each reflector, which can be as high as 98%, as well as on the separation between reflectors. The reflectors may consist of a periodic array of overlaid stripes, either metallic (field-shorting type) or dielectric (mass-loading type), or an array of etched grooves as in the reflective array structures.

When an RF voltage of the proper frequency is applied to the IDT, SAW's are generated and propagate away from the transducer in both directions. The overlap of the IDT is large compared to the acoustic wavelength to minimize diffraction, and the acoustic waves have constant phase across the entire transducer aperture. As the waves enter the grating reflector, coherent multiple reflections occur and a large standing wave is set up in the cavity. The standing wave produces a profound effect upon IDT impedance characteristics, resulting in a high quality factor ($\approx 10^4$) at resonance. SAW cavities with resonant frequencies as high as 1 GHz have been made.

Another type of SAW resonant cavity consists of two IDT's with a single grating reflector placed be-

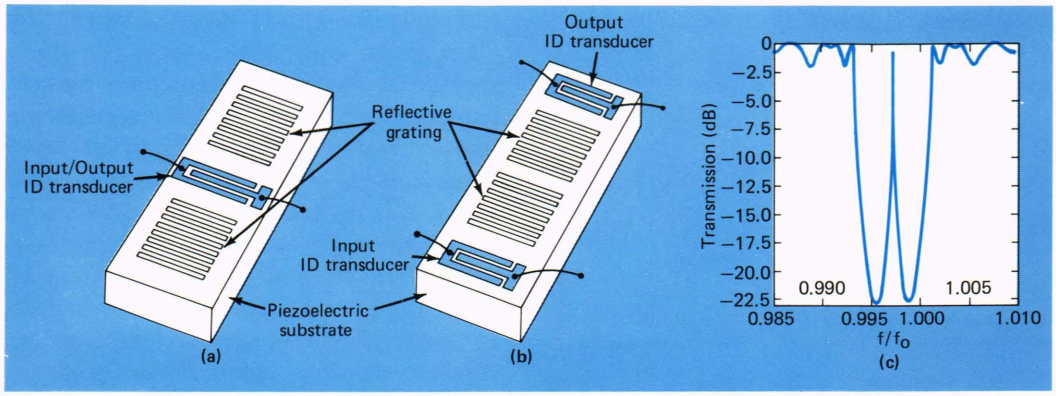


Fig. 12—SAW resonant cavity structures and a typical response function.
(a) Single transducer inside two reflective gratings.
(b) Two reflective gratings inside IDT's of basic delay line forming resonant cavity.
(c) Frequency response of resonant cavity (b).
(Q is typically 6.15×10^3, f_0 equals 68.5 MHz)

tween them. The two IDT's form a surface wave delay line with a transfer function proportional to $[\sin(2\pi f)/2\pi f]^2$ and a large fractional bandwidth if each IDT is composed of a few finger pairs. The SAW reflector is composed of a large number of reflector sections that consist of a quarter wavelength section of impedance Z_1 (the strip) and a quarter wavelength section of impedance Z_0 (the gap). Because of the impedance discontinuity, each reflective element contributes a small reflection of the SAW as it propagates through the section. The bandwidth over which the surface waves are reflected is much smaller than the transducer bandwidth if the number of reflecting sections is large compared to the number of IDT finger pairs. The reflector thus creates a narrow stopband notch in the passband response of the delay line.

The two-transducer configuration may be used as a cavity resonator by placing two of these gratings between the transducers to reflect the surface wave upon itself many times. The structure is shown in Fig. 12b; a typical response curve is given in Fig. 12c.

Multistrip Coupler. The multistrip coupler employs the second-order effects to couple energy from one area of a SAW medium to another.[5] This coupling is the result of the alternating potential difference between metal strips and the acoustoelectric wave beneath the fingers of each ID comb (Fig. 4b). The metal fingers conduct these potentials along their length to generate new acoustic waves. Multistrip couplers have been used for track changes, reflectors, unidirectional transducers, and reflecting track changers.

Oscillators/Waveform Generators

Oscillators. Lewis has shown that a transistor oscillator with a SAW filter in the feedback loop (Fig. 13a) will exhibit great stability.[13] The SAW filter characteristic controls the frequency of oscillation. If it is built on a crystal that has temperature compensation, the frequency stability will be comparable to that of conventional quartz crystal oscillators. Both aluminum nitride piezoelectric films and silicon layers can be deposited on sapphire, thus giving rise to the hope that this oscillator can be fabricated as an integrated component.[14]

[13] M. F. Lewis, *Ultrasonics* **12,** 1974, 115.

[14] P. J. Hagon, F. B. Micheletti, R. N. Seymour, and C. Y. Wrigley, *Trans. Microwave Theory Tech.* **MTT-21,** 1973, 303.

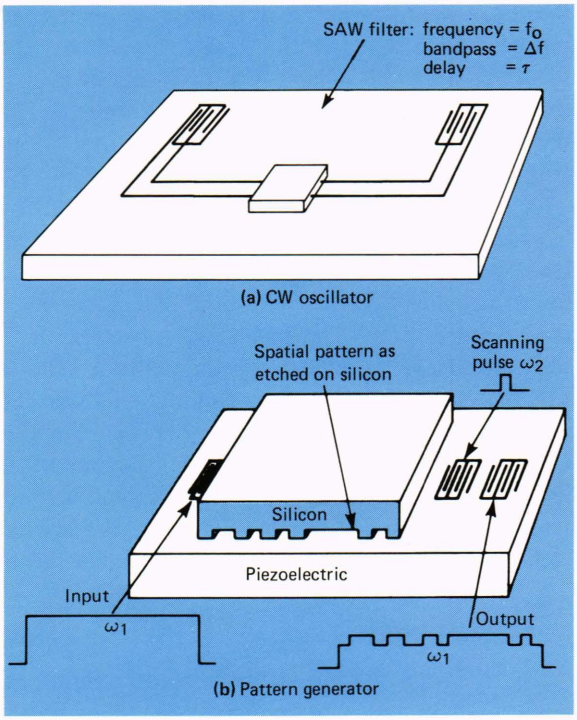


Fig. 13—SAW oscillators are formed by placing SAW filters in the feedback loops of transistor amplifiers. Nonlinear elements and overlays give rise to pulsed signals and pattern generation.

Waveform Generators. Many new SAW devices exist for transforming various signals and patterns into such forms as convolvers, Fourier transformers, optical scanners, and acousto-optic devices. Each must be driven by a special waveform. Typically, the convolver requires a short pulse for the time reversal of the signal. In pattern recognition, a wide variety of waveforms is required. Fourier transforms and various acoustic optic devices require a linear chirp as the driving source.

Long pulses (duration greater than the SAW filter delay) may be generated by using the Lewis oscillator (mentioned above) simply by turning the amplifier off and on (i.e., by applying a rectangular waveform to the power supply). For short pulses, this technique does not work so a new, regenerative pulse amplifier (RPA), method is used. In the RPA method, a nonlinear element is included in the SAW oscillator feedback loop. A silicon epitaxial layer adjacent to the SAW filter could be used. This layer attenuates (absorbs) the acoustic energy, but it does saturate; under these conditions high-level waves suffer less attenuation than low-level waves. In this case, the oscil-

lator will act as an RGA and generate a series of pulses whose bandwidth is determined by the system bandwidth (the SAW filter bandwidth if the transistor amplifier bandwidth is large).[15] The repetition rate is determined by the total loop-transit time.

A chirp generator could be constructed by using a dispersive filter driven by the short pulses from the RGA. Pattern generation can be achieved by using the electric field coupling between a silicon layer and a piezoelectric substrate. It is only necessary to etch a spatial pattern onto a silicon eptaxial layer. The etching changes the spacing and hence the coupling between the silicon and the piezoelectric. When this component is driven with both an RF signal pulse and a short scanning pulse as shown in Fig. 13b, the spatial etched pattern is transferred onto the signal pulse.

Matched Filters

In signal processing of radar and communications systems, it is often necessary to extract a signal pulse in the presence of noise, e.g., to detect a peak power-limited radar signal. Correlating the active multiplication of a signal and a reference waveform followed by summation or integration techniques can perform the required detection in the presence of noise when the shape of the input signal is known. The process involves the convolution of the input signal, $V_{in}(t)$, with a weighted detection function, $h(t)$, such that the output signal is given by

$$V_{out}(t) = \int_{-\infty}^{\infty} h(t)V_{in}(t - \tau)\, d\tau.$$

The optimum weighted detector function to use for improving the output signal-to-noise ratio in the presence of band-limited white Gaussian noise is the time inverse of the input signal. Thus, the detector is "matched" to the input signal and is called a matched filter. The term usually has been applied to a tapped delay line or nonrecursive filter structure (passive device). SAW devices can be used to implement efficiently either a matched filter or correlator.

SAW versatility allows a large number of waveforms to be generated and match-filtered, including linear chirps and a host of non-FM waveforms such as Barker codes, biphase codes, higher polyphase codes, burst processing sequences, and pseudorandom codes.

Figure 14a illustrates a generated and correlated fixed code signal. Any number of codes, from a few-bit Barker code to a thousand-bit pseudorandom sequence, can be implemented in this manner. For a reasonable number of bits (≤ 100) at a frequency of 100 MHz, the design and fabrication are relatively straightforward. As the number of bits increases, the number of ID fingers increases, and problems of second-order effects come into play. Ways to compensate for these problems include the use of split electrodes, dummy fingers, dielectric film overlays, and tilted transducer geometry.[5]

A simple, nonperiodic, coded burst of RF energy (Fig. 14b) can be used as an effective

[15] C. Culter, *Proc. IRE* **43,** 1955, 140.

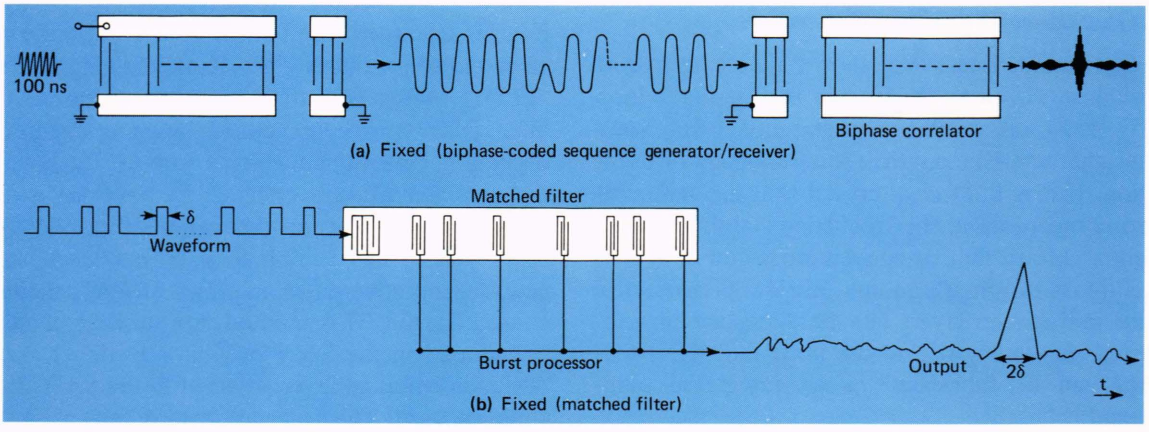


Fig. 14—SAW coded devices.
(a) Fixed (biphase coded sequence generator/receiver).
(b) Fixed (matched filter).

radar signal. If the burst is processed through the matched filter shown (a tapped delay line), and if the original burst were carefully designed, the output is one main pulse with very low sidelobes. If the code has n pulses arranged so that no more than two pulses overlap upon correlation (except when they all overlap) the peak-to-sidelobe ratio is $n/1$. The sidelobes stretch out on either side of the correlation peak for a time equal to the pulse length.

Programmable "matched" SAW filters are needed, especially for identification, synchronization, and spread spectrum communications. Although these SAW devices are not as refined as the fixed-code matched filters, their feasibility has been demonstrated and they will be available for future systems application.

Convolvers

The SAW devices discussed so far operate below the saturation level of the acoustic medium. The medium becomes nonlinear at a point where Hooke's law is not obeyed (i.e., stress is not proportional to strain). These effects become noticeable when relative strain reaches levels on the order of 10^{-4}.[10] This nonlinear behavior leads to parametric interactions between different RF signals.

Figure 15a represents a degenerate SAW convolver. When applied to the IDT's, the two input waveforms launch surface waves toward the center of the device. The substrate is required to give a nonlinear relationship between the electric field and the acoustic strain, such that fields proportional to the product of the two SAW strains are produced. The fields are detected by the parametric electrode. The term "degenerate" refers to the fact that the interaction is strongest when the two input waveforms have the same center frequency.

Although the interaction mechanism is not fully understood,[16] the important features may be explained by the fact that in suitable materials the nonlinear piezoelectric relation can be expressed as

$$D = \epsilon E - eS + KS^2,$$

where D is the electric displacement, ϵ is the dielectric constant, E is the electric field, e is a piezoelectric constant, and S is the acoustic strain. Constant K, which determines the nonlinear interaction, is assumed small so that SAW strain can be regarded as undisturbed by the interaction.

Assume two counter traveling waves at the same frequency with strains

$$S_1 = a_1 \cos(\omega t - kx)$$

and

$$S_2 = a_2 \cos(\omega t + kx).$$

The total strain, $S = S_1 + S_2$, will produce a term in the above electric displacement equation proportional to $a_1 a_2 \cos 2\omega t$. This term has frequency 2ω and no spatial dependence and is detected by the parametric electrode (a uniform metal film acting in conjunction with a ground plane on the

[16] F. G. Lean and C. C. Tseny, *J. Appl. Phys.* **41**, 1970, 8912.

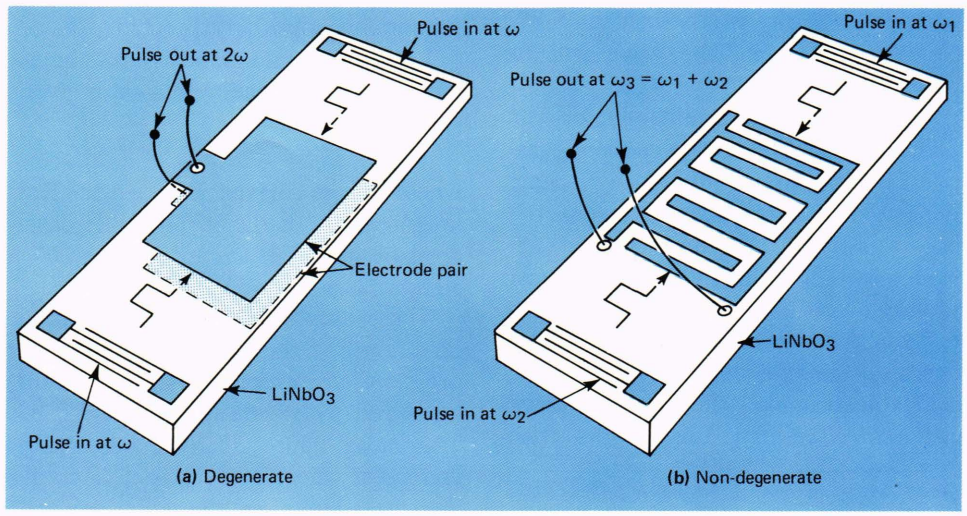


Fig. 15—SAW convolvers/correlators.

reverse side of the substrate).

The detection by a plate of length, L, implies that the signal at 2ω is integrated over L. It can be readily shown that the detected signal is the convolution of the two input signals.

Nondegenerate convolvers (Fig. 15b) are designed to operate with different center frequencies, ω_1 and ω_2, at the two input ports. It can be shown that the product term at frequency ω_3 has the form $a_1a_2 \cos [\omega_3 t + (k_2 - k_1)x]$, where $\omega_3 = \omega_1 + \omega_2$ and k_1 and k_2 are the wave numbers of the input SAW's. The associated electric fields can be detected by an IDT whose periodicity corresponds to the spatial variation of the product term, i.e., an IDT synchronous at $\omega_2 - \omega_1$. Note that this does not correspond to the output frequency and that the expression for the product term does not imply excitation of a SAW at frequency ω_3.

Amplifiers

SAW amplifiers come in two basic types: combined media amplifiers (CMA)[17] and separated media amplifiers (SMA).[18] A typical CMA is shown in Fig. 16a. Cadmium sulfide (CdS), which is piezoelectric and photoconductive, is used for the substrate. Under illumination, carriers that are produced near the surface interact with the electric field of a Rayleigh wave propagating along the surface. If a voltage is applied across the illuminated region, the carriers can be caused to drift in the direction of the propagating surface wave. If the carrier velocity is greater than the SAW velocity, energy will be delivered to the wave, and its amplitude will increase with distance. If the carrier velocity is less than the SAW or in the opposite direction, the wave will be attenuated.

In SMA, the functions of semiconductor and acoustic media are separated. A typical configuration, shown in Fig. 16b, uses a thin film ($\approx$1 μm) of silicon deposited on sapphire as the semiconductor; this is separated from the acoustic substrate ($LiNbO_3$) by silicon dioxide (SiO_2) spacer rails or randomly placed dots like the pilings on a fishing pier. Typical spacings are of the order of 50 nm. There is a gain, just as in the CMA's, when the drift voltage makes the carrier velocity greater than the SAW velocity. Loss occurs when

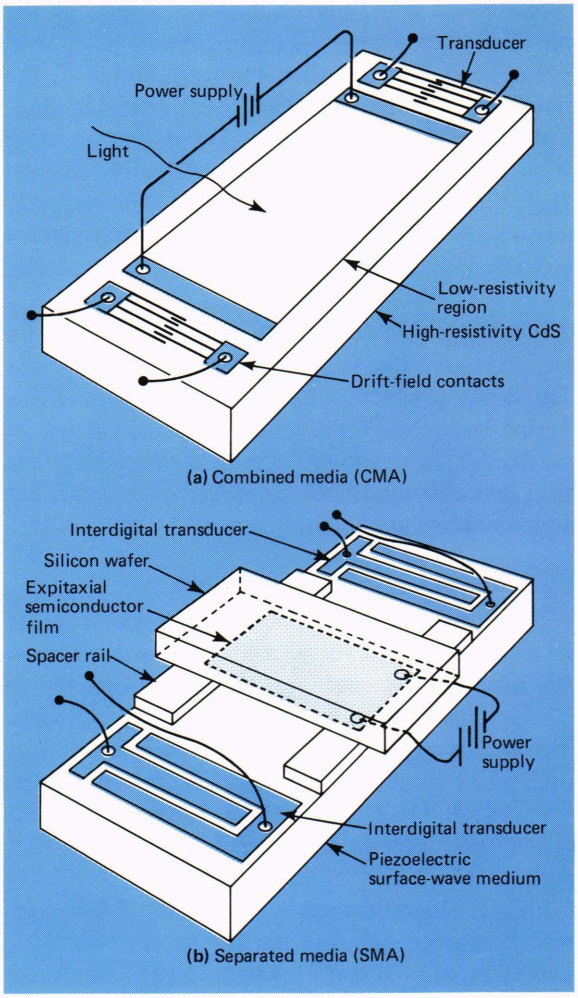


Fig. 16—SAW amplifier structures.

the velocity is less than that of the SAW or the potential is oppositely directed.

Frequency gain depends on the gap spacing. The spacing for maximum gain is given by

$$2\pi h/\lambda_R = \epsilon_O/\epsilon_P,$$

where h is the gap spacing, λ_R is the Rayleigh wavelength, and ϵ_P is the effective dielectric constant of acoustic material. On $LiNbO_3$ at 200 MHz (ϵ_P = 50 ϵ_O), h is 50 nm for maximum gain.

These amplifiers exhibit loss in the backward direction and thus give good triple transient suppression. Noise figures as low as 5 dB have been measured with saturation levels as high as 1 W. The noise figure decreases as the gap spacing increases, whereas the power handling capability increases. However, the gain decreases. Projected efficiencies may eventually reach 10% with low

[17] R. M. White and F. W. Voltmer, *Appl. Phys. Lett.* **8,** 1966, 40.

[18] J. H. Collins, K. M. Kakin, C. F. Quate, and H. J. Shaw, *Appl. Phys. Lett.* **13,** 1968, 314.

noise and operating frequencies from 50 to 1000 MHz.

Phase Shifters

The use of an external electric or magnetic field applied to the propagation path of a SAW (e.g., between the two transducers of a SAW delay line) will produce a change in SAW velocity, thus giving rise to a phase shift. The external field is usually applied through an interaction medium (thin-film overlay). Electric field phase shifting is accomplished by using an electrostrictive overlay or by direct coupling in piezoelectric materials.

A basic magnetic SAW phase shifter[19] consists of a magnetostrictive thin film (e.g., 0.85-μm nickel, vacuum-evaporated) deposited on a SAW delay line between the two IDT's. The acoustic wave is launched in a Rayleigh mode but is converted to a layer mode as it passes under the film. Since the film is magnetostrictive, its effective elastic moduli depend on the direct magnetic bias field applied. Any change in the elastic moduli in the film changes the velocity of the acoustic wave propagating in the layered media and hence produces a phase shift. In typical experiments, phase shifts as high as 80° at 210 MHz have been reported.

Acousto-Optics

The propagation of an acoustic wave in the bulk or on the surface of a transparent material produces a periodic modulation of the index of refraction by means of the elasto-optic or photoelastic effect. The resulting "optical phase grating" may diffract portions of an incident light beam in one or more directions. The grating spacing is equal to the acoustic wavelength and may be varied by changing the frequency of the acoustic signal. The acoustically established and controlled phase grating may be used for the deflection's modulation and filtering of an incident light beam, thus allowing the implementation of devices such as optical switches, modulators, correlators, scanners, turnable filters, and spectrometers. Planar acoustic optic devices using SAW offer increased performance, reduced size, reduced power consumption, and better control and phase matching than their bulk wave counterparts.[20]

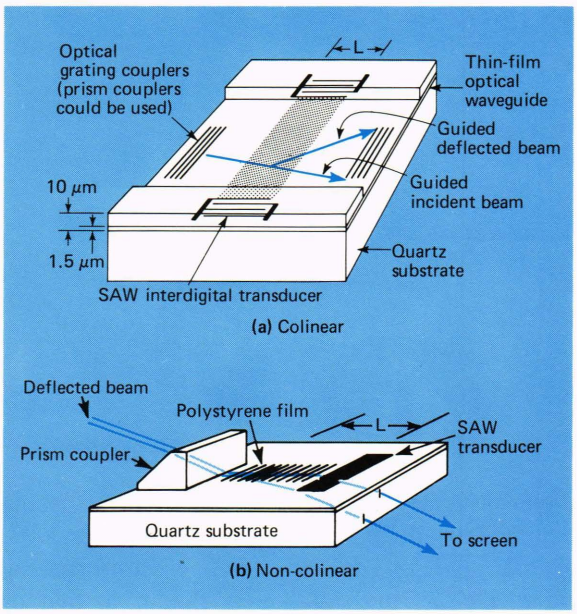


Fig. 17—Acousto-optic deflectors.

Acousto-Optic Deflectors. There are two types of acousto-optic deflectors: (a) the incident and deflected light beams are both guided by optical waveguides as shown in Fig. 17a, and (b) the SAW deflects the incident guided light beam out of the waveguide into an unguided radiation mode (Fig. 17b). In the first type, Bragg efficiencies up to 93% have been reported,[21] while in the second type the angle at which the beam escapes may be changed by varying the SAW frequency.[22]

Acousto-Optic Mode Converters. An acousto-optic mode converter is essentially an acousto-optic deflector in which a thin-film optical guided wave is diffracted from one mode to another by a SAW.[23] Figure 18 illustrates a typical experimental device. The waveguide is 1.6-μm-thick electron-beam-evaporated 7059 glass film on an aluminum-coated $LiNbO_3$ substrate. The laser beam ($\lambda = 632.8$ nm) was coupled into and out of the waveguide with grating couplers. The SAW is launched by an IDT parallel to the guided light beam. At 320 MHz (SAW) with proper phase adjustment, there was a 55% coupling from the TE_3 to the TE_1 waveguide mode.

Acousto-Optic Modulators. Modulation occurs due to diffraction of the guided optical wave by

[19] A. K. Ganguly, K. L. Davis, D. C. Webb, C. Vittoria, and D. W. Forester, *Electron. Lett.* **11,** 1975, 610.

[20] R. V. Schmidt, *IEEE Trans. Sonics and Ultrasonics* **SU-23,** 1976, 22.

[21] K. W. Loh, W. S. C. Chang, and R. A. Becker, *IEEE J. Quantum Electron.* **QE-11,** 1975, 86D.

[22] F. R. Gfeller and C. W. Pih, *Electron. Lett.* **8,** 1972, 549.

[23] L. Kuhn, P. F. Heidrich, and E. G. Lean, *Appl. Phys. Lett.* **19,** 1971, 428.

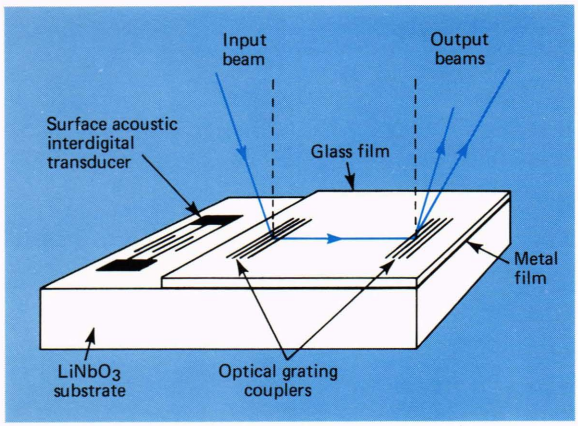


Fig. 18—Acousto-optic mode converter.

the acoustically established phase grating. Specific details depend upon grating width (W) and the grating spacing (d). For large width-to-spacing ratios (W/d), the situation is similar to that of X-ray (Bragg) diffraction by a crystal lattice. Conservation of momentum leads to the Bragg condition for beam deflection single (θ_B). This type of modulator, which is also a switch, has the advantage of giving amplitude modulation without the need for an additional polarizer as is usually required. The intensities of the undiffracted and diffracted beams are $I_o\ (\cos^2 \phi/2)$ and $I_o\ (\sin^2 \phi/2)$, respectively, where ϕ is the acousto-optically induced phase shift.

Acousto-Optic Correlators. A SAW properly coupled with guided optical waves may be used to produce convolution (correlation) in thin-film optical waveguides and potentially in single-mode optical fibers. In a typical optical guided wave-SAW convolver,[21] the optical guided waves from two SAW transducers are appropriately combined to give an electrical output signal from an optical heterodyne receiver, which is the convolution of the input signals.

Chirp Transformation

Chirp transformation (CT) is a signal-processing technique for obtaining the Fourier transform (FT) of a signal by mapping the frequency domain into the time domain.[24] In its most common form, the CT consists of premultiplication of the signal by a waveform whose frequency varies linearly with time (chirp), followed by a convolution in a chirp filter with opposite frequency-modulation slope. The FT is obtained by a final multiplication with a chirp identical to the premultiplication stage. This method is termed MCM (multiplication-convolution-multiplication) processing. A discrete form of CT, known as chirp z-transform (CZT), was developed by Rabiner et al.[25] for the computation of the discrete FT (DFT) of sampled data. CZT techniques are now routine in digital processing.

Although CT methods have been known for several years, it was not until the advent of SAW product chirp filters with high time-bandwidth products that real-time CT processing of analog signals could be considered. Charge-coupled devices (CCD's)[26] have also provided powerful means of implementing chirp filters. However, CCD filters operate from baseband up to approximately 10 MHz (experimentally $\simeq$ 100 MHz), while SAW chirp filters cover the range from 10 MHz to 1 GHz. Thus CCD and SAW appear to be complementary technologies for analog CT processing.

Several real-time signal-processing techniques based on analog CT processing are possible including spectral analysis, network analysis (determination of network transfer function), and signal filtering (linear filtering-direct synthesis of transfer function), with the possibility of externally programming the filter function, signal generation, and synthesis.

6. SAW at APL

The Advanced Technology Section of the Microelectronics Group acts as a center for in-house SAW awareness and in this capacity has made many contributions to Laboratory projects. Several prototype devices, primarily filters and delay lines, have been made and tested. In addition, an *in situ* thin-film vacuum deposition rate and thickness monitor was developed that directly incorporates SAW phenomena.

The Group entered SAW technology approximately five years ago, at a time when microelectronic techniques were used to fabricate transducers for research studies. These investigations used time and space resolution in mapping the radiation of bulk waves concomitantly generated with the SAW. SAW frequencies up to 250 MHz

[24] A. Papoulis, *Systems and Transforms with Applications in Optics*, McGraw-Hill, NY, 1968.

[25] L. R. Rabiner et al., *Bell Syst. Tech. J.* **48,** 1969, 1249.

[26] H. J. Whitehouse, *IEEE Trans. Solid State Circuits* **SC-11,** 1976, 64.

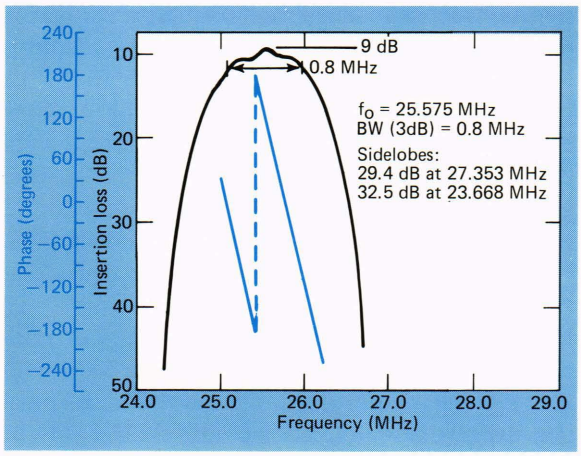


Fig. 19—Performance characteristics of slit-slit transducer filter (insertion loss and phase shift versus frequency).

were required, and photolithographic fabrication methods with resolutions approaching 1 μm had to be developed. Using these acquired fabrication techniques, the Section designed several delay lines and bandpass filters. While the first devices were of a basic nature, they aided our understanding of SAW devices and gave insight into their potential applicability. The performance characteristics of one of these filters is shown in Fig. 19. It is a 25-MHz two-transducer filter on a lithium niobate substrate. Each transducer consists of 35 finger pairs with a slit-type geometry (more metal than gap). Slit-type transducers reduce spurious responses due to reflected signals and regeneration. The salient filter parameters are 9-dB insertion loss (no external matching networks), 0.8-MHz bandwidth, 25.575-MHz center frequency, 1.5-dB ripple, 20-dB sidelobe rejection, and better than 1° phase linearity. The filter could be packaged in a 0.5-cm × 0.5-cm × 0.2-cm volume.

Other areas of interest are (a) mass loading effects (change in velocity, phase, or mode of a signal transversing a region coated with a deposited material(s)), (b) velocity perturbations caused by applied magnetic and electric fields, and (c) the influence of SAW on the adhesion of deposited films. The change of velocity with applied magnetic field has been used to make variable phase shifters for communications subsystems. We plan to incorporate this effect into similar components for evaluation. The authors have observed that surface waves propagating in a substrate during deposition result in increased film adhesion to that substrate; thus it might be possible to adhere normally incompatible materials. Little has been done in this area, but it suggests a phenomenon for future investigation.

The mass loading effect has been applied by the authors in making a thin-film deposition rate and thickness monitor. It consists of a 30-MHz two-transducer SAW delay line on a temperature-stable quartz substrate (ST cut). The monitor package has a series of 12 apertures exposing segments of the path between the two transducers to an evaporant stream. For maximum sensitivity and range, a different aperture (virgin underlying surface) is selected for each deposit; the same aperture can be used again if excessive mass (<8 μm of aluminum equivalent) has not accumulated. An exploded view of the monitor head is shown in Fig. 20.

This instrument measures changes in surface velocity due to free surface loading. It is a linear phenomenon proportional to the total mass per

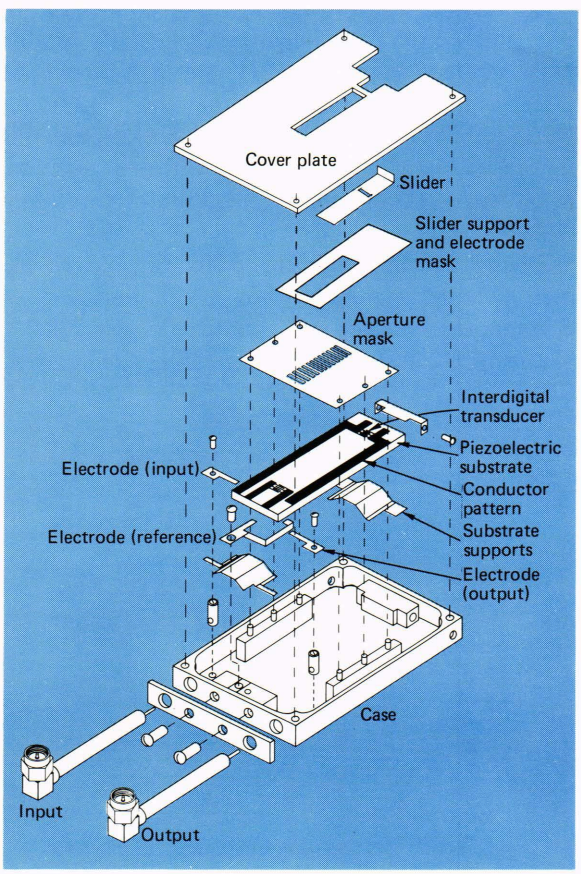


Fig. 20—SAW thickness and rate monitor assembly drawing.

unit area of the film layer. The resulting phase difference, ϕ_d, between a "free" wave and "loaded" wave traversing a region of length l is

$$\phi_d = \frac{2\pi f l}{V_f}(1 - V_f/V_L),$$

where V_f is the free surface velocity, V_L is the loaded surface velocity, and f is the operating frequency.

The phase shift due to mass loading is monitored by comparing, by means of a phase detection or vector voltmeter, the phase difference between two parallel paths, one exposed to the evaporant stream and the other not. Figure 21 shows a typical response using chromium. When depositing conductors on the monitor's piezoelectric sensor, there is a large initial phase change due to the shorting of the surface fields. The subsequent phase changes are then due to the mass loading effect. The trace in Fig. 21 shows that, of the approximately 14° of total phase shift, 1.5° is associated with surface shorting. Using the total film thickness (determined by an optical interferometer) of 289 nm, we obtain a linearized sensitivity for the mass loaded region of 23.1 nm per degree of phase shift. Since the phase can be read to at least ±0.2°, a conservative resolution for this device (i.e., at $f = 30$ MHz with an aperture 0.127 cm × 0.406 cm) of ±4.5 nm of chromium is predicted.

Figure 22 illustrates the effect of gold as the mass loading material. With a film thickness of 570 nm, the total recorded phase shift was 102.3°. The initial offset was 1.7°, yielding a linearized

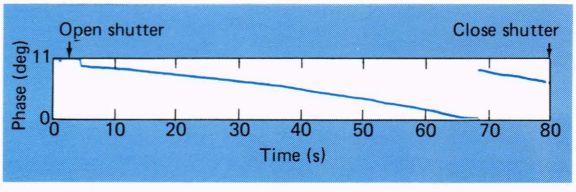


Fig. 21—SAW monitor chromium evaporation (phase shift versus time).

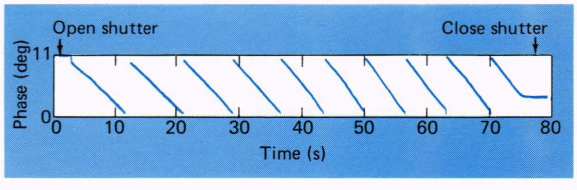


Fig. 22—SAW monitor gold evaporation (phase shift versus time).

sensitivity of 5.7 nm per degree and a resolution of ±1.0 nm with gold. The sensitivity for gold is four times as great as that for chromium (the phase shift is four times as great for a given thickness of gold as for the same thickness of chromium). Since the density of gold is 19.3 gm/cm^3 and the density of chromium is 7.2 gm/cm^3, this suggests a sensitivity ratio of 2.68. The discrepancy may be resolved by considerations relating to the experimental conditions. However, since repeatability and proof-of-principle were of primary interest, we stressed qualitative rather than quantitative behavior. In passing, the resolution of this monitor could easily be increased by an order of magnitude by increasing both the frequency of operation and mass-loaded pathlength.

A particularly useful property of the SAW thickness monitor is its ability to record effective (average) thickness before formation of a continuous film. This should prove helpful to those investigating the physics of film nucleation.

Future plans are to integrate the frequency source and phase detector so that only a DC power supply and DC readout/recorder would be required for operation.

7. Summary

In summary, the simplified analytical descriptions along with the detailed applications section should provide insight into the design and structuring of useful SAW devices. It is evident that, while SAW technology is conceptually simple, it is mathematically complex, especially if known higher order effects are to be incorporated. Computer design is required for all but the most elementary devices. Although the range of applications for SAW technology is quite large, with new areas being continually developed (e.g., APL's SAW thickness monitor), it is not all encompassing. In fact SAW technology has steadily progressed (determined by systems applications and field use) only in such areas as dispersive delay lines, bandpass filters, and tapped delay lines. Part of this restricted use stems from the lack of understanding of SAW technology by systems designers and part from technology limitations. We trust that this article has created a basic understanding of microsonics and provided the system designer with another useful tool in his "bag of tricks."