AN OVERVIEW OF ACOUSTIC SURFACE-WAVE TECHNOLOGY

G.W. Farnell and E.L. Adler

AN OVERVIEW OF ACOUSTIC SURFACE-WAVE TECHNOLOGY

G.W. Farnell and E.L. Adler

AN OVERVIEW OF ACOUSTIC SURFACE-WAVE TECHNOLOGY

G.W. FARNELL and E.L. ADLER

Department of Electrical Engineering
McGill University

Final Report to
Communications Research Centre
on DSS Contract 36001 - 3 - 4406

Contract Supervisors

D.B. Wohlberg

S.M. Chow

August 12th, 1974

ABSTRACT

This report presents an overview of the principles of surface-wave devices important in signal-processing applications. The fundamental properties of surface-waves are reviewed and currently practical components for use in radar systems, spread-spectrum communications systems, and bandpass filtering are described. The main emphasis is on dispersive and non-dispersive delay lines, devices for generating and detecting various waveforms, and methods for generating and detecting fixed and programmable biphase codes. The discussion of new devices includes multistrip couplers, oscillators, long delay lines, surface-wave convolvers and amplifiers, reflective array structures, and devices using overlay films. The properties of substrate materials and current fabrication methods of importance in device design are covered in considerable detail. Performance specifications of current operational components are included throughout the report.

It is concluded that surface-wave devices have already found many applications because of their size, cost-effectiveness, and reliability; larger scale applications will come as systems engineers become sufficiently familiar with these devices and their potential to incorporate them into operational designs.

TABLE OF CONTENTS

I. INTRODUCTION

The intent of this report is to provide the radar and communication system designer with an overview of the fundamental operating mechanisms of surface-wave devices, their realizable operating characteristics and their limitations. It appears that surface-wave devices will play a small but important role in various parts of a variety of communication systems, but that these devices will usually be integrated into subsystems rather than purchased as a line of discrete components. Thus the non-specialist in surface-wave components is likely to be more involved in the understanding of their principles than he would be if he were selecting pre-packaged standard components. To this end the report tries to discuss most of the aspects of surface-wave technology which are of current importance or potential usefulness. The discussion is rather cursory in many places, of necessity because of space restrictions but more important because of the philosophy of attempting only to provide the systems engineer with a conceptual outline of the significant ideas and order of magnitude estimates of the currently available performance parameters along with a reasonable bibliography from which the detailed characteristics can be sought.

The authors realize that it is very unwise for engineers concerned with components to presume that every new device has a system application and we have tried to avoid this "solutions - seeking - problems" syndrome and restrict the discussion to concepts and components which have at least been considered for system use.

Outline of Report

As noted before, this report is intended as an overview of the principles of surface-wave devices and of the realizable characteristics useful for system applications. While it is desirable for continuity to read through the chapters successively, the reader will find that an attempt has been made to keep the chapters somewhat independent so that he need only consult those sections of interest to him.

References from which more details concerning particular components can be obtained are sprinkled liberally through the text and in Appendix I a short guide to the literature is provided which indicates the key papers in the different areas. The bibliography is quite extensive, but it was decided to restrict it to papers and reports of direct relevance to material in this report.

The central concepts for device applications of surface waves are contained in Chapters II and III which are concerned with the principles of propagation, generation and detection. These principles are then applied to the production of a range of terminal characteristics. The currently operational devices are described in Chapters IV to VII which treat linear-phase band-pass filters, large time-band width product chirp filters of two types, and phase-code generators and correlators for pulse-compression radars and spread-spectrum communications. Chapter X outlines the important characteristics of the available materials for surface-wave substrates and the fabrication procedures for devices. The remaining chapters cover some important features available in surface-wave technology which are of at least potential system usefulness.

To place surface-wave components in their current system context we take the liberty in Appendix III of quoting verbatim the interim conclusions from a report made in February 1974 by Group KAG 3 of The Technical Cooperation Program (TTCP) of their study entitled "The Application of Surface Acoustic Wave Devices to Radar Signal Processing". While radar applications are emphasized, many of the comments could be carried over directly to other forms of communication systems.

The performance specifications of many devices are dispersed through the text but for general illustrative purposes a few examples of operational components have been added in Appendix II.

II. FUNDAMENTALS OF ACOUSTIC SURFACE WAVES

The physical phenomenon which is the central operating mechanism common to all of the signal-processing devices to be considered in this report is the ability of the free surface of any solid to support a mode of elastic (mechanical) wave propagation confined to the region of the solid immediately below the free surface. Thus we will be dealing with propagating ripples on the solid surface as indicated schematically in Fig. 2.1.

The properties of these waves will be considered in some detail in various sections, but the important characteristics can be summarized here. First the wave carries vibrational energy parallel to the surface and the frequency range of interest is in the mega- to gigahertz range, thus it will be called an acoustic surface-wave* and while generalizations will appear later, the particular acoustic surface-wave of most importance is also called a Rayleigh wave (G1, Chapt. X). The velocity of propagation is in the range from 1,000 to 6,000 m/sec., and independent of frequency, thus the wave-length is of the order of 30 microns at 100 MHz (G2). If the solid on which the wave is propagating is piezoelectric then an electric field propagating with the wave extends into the space above the solid and the interaction of this field with metal electrodes on the surface can be used to excite the wave, to tap it, and modify it as it progresses. For chosen solids the attenuation of the surface wave can be negligible up to frequencies of many hundred megahertz.

---

* The acronym SAW (Surface Acoustic Wave) has been coined for these waves but will not be used extensively here.

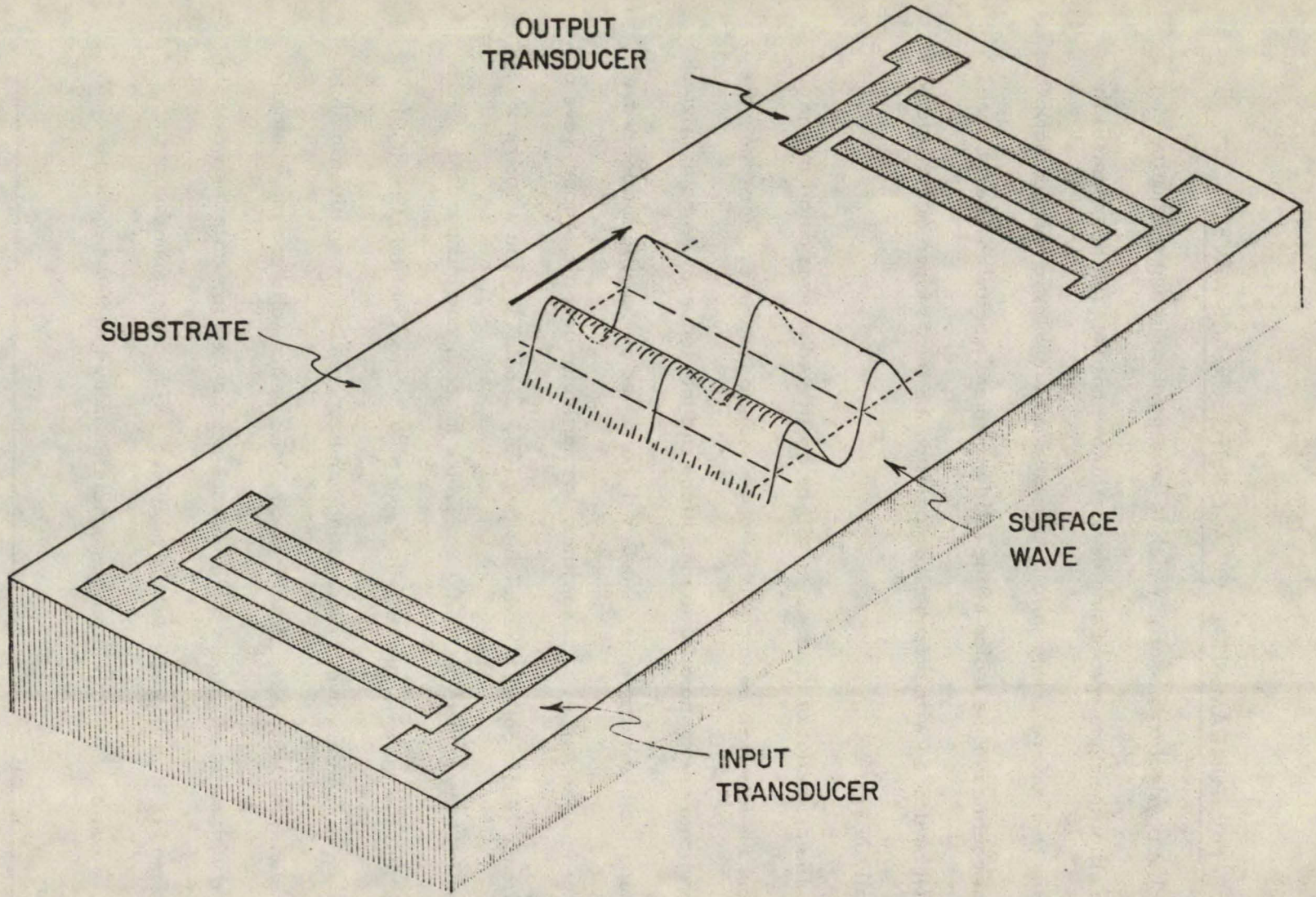


Fig. 2.1 Prototype geometry for surface-wave device.

Because of these properties of slow dispersionless velocity, low propagation loss, energy transport close to the surface and convenient interaction with external circuits, acoustic surface-waves form the basis of a useful set of signal processing devices including delay lines, band-pass and transversal filters, chirp filters, phase-code generators and correlators, oscillators, signal transformers, convolvers and several other potentially important components (G 9). The electrode and coupling structures which are fabricated by standard photolithographic techniques can be quite complicated to provide system transfer functions, difficult to realize by other means, in devices that are miniature in size and compatible with the planar technology of other integrated circuit components.

2.1 Velocity and Displacement

Before looking at the Rayleigh wave in more quantitative detail, let us note other modes of acoustic wave propagation which help the understanding of surface-waves but which to the device designer are usually an undesirable source of spurious signals. In an unbounded isotropic solid three mechanical plane waves can propagate in any direction, one polarized along the propagation vector and called a longitudinal bulk wave and two (here degenerate) polarized perpendicular to the propagation vector and called shear or transverse bulk waves. In a bounded solid these modes can be reflected, mode coupled or excited at the bounding surfaces. The shear wave velocity v_t is less than that of the longitudinal wave v_l and of course in an isotropic solid both are independent of direction.

Now, continuing with isotropic solids, the velocity of propagation v of the acoustic surface-wave on a free plane surface is somewhat less than the shear

velocity as indicated for the complete range of possible ratios in Fig.2.2 (G 3). Due to its lower velocity, the surface-wave propagating on an unperturbed free surface cannot phase match and thus couple to a bulk mode propagating at any angle to the surface. However such coupling will be seen to occur if the surface is perturbed.

The particle displacement of a Rayleigh wave has two components, one parallel to the propagation vector and one normal to it and to the surface. The two components are in phase quadrature and their separate variations with depth are shown in Fig.2.3. It is noted that the amplitude of the particle displacement becomes negligible for depths greater than a few wavelengths and recalling that if the frequency is 100 MHz and the velocity 3500 m/sec this wavelength is but 35 microns.

The form of the displacement of the solid near the surface is better shown in Fig.2.4 which illustrates the instantaneous distortion of a square grid in a crosssectional or sagittal plane. The retrograde elliptical motion of a point on the surface is here clearly evident. The amplitude of the displacements shown is greatly exaggerated, in typical signal processing applications displacement amplitudes are less than $10^{-5}\lambda$ for the milliwatt signal levels used. There is an acoustical analogue to the Poynting vector of electromagnetic waves giving the power flow across a unit area perpendicular to the direction of propagation (G 3, Chapt.5). Figure 2.5 shows the variation with depth of such an acoustic Poynting vector (66) and one again notes that all of the energy in the surface-wave is carried close and parallel to the surface, the depth of penetration being only of the order of a wavelength.

The diagrams for Figures 2.2 to 2.5 were drawn for an isotropic solid as substrate. In signal processing applications the frequencies involved are usually so high

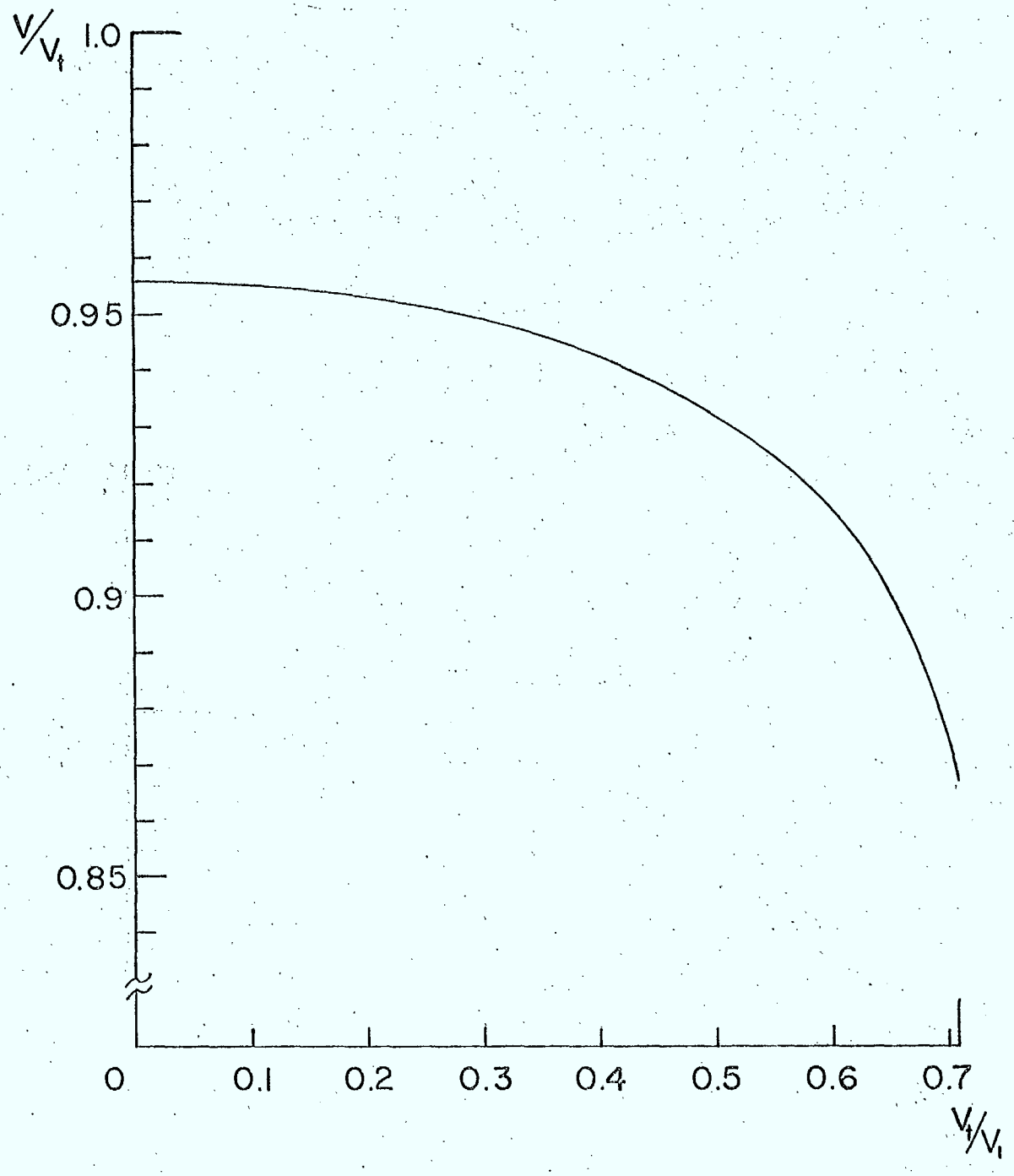


Fig. 2.2 Surface-wave velocity on isotropic substrates. v_t = transverse or shear bulk-wave velocity. v_l = longitudinal bulk-wave velocity.

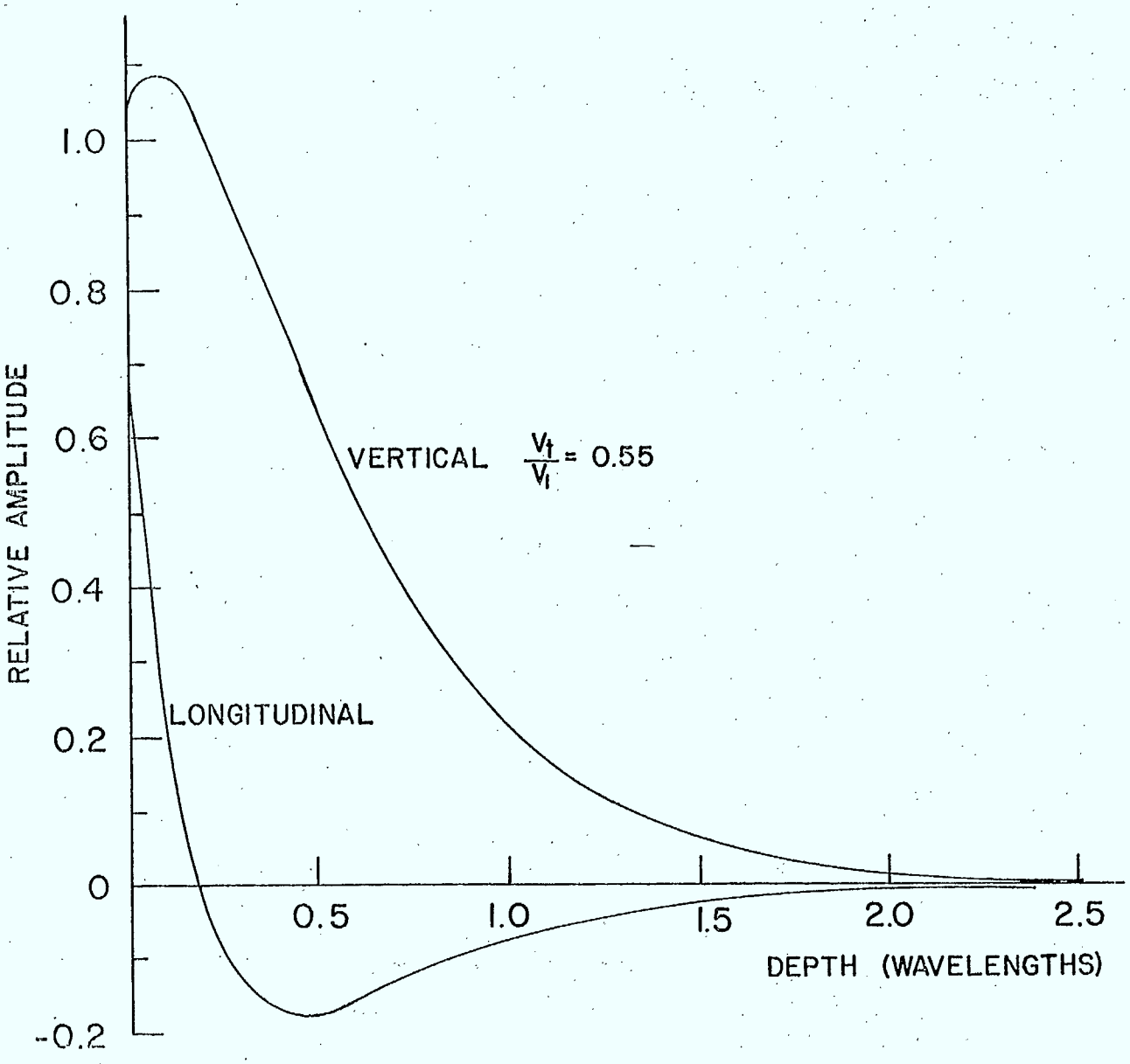


Fig. 2.3 Amplitude of displacement components as a function of depth. Displacement components are in phase quadrature.

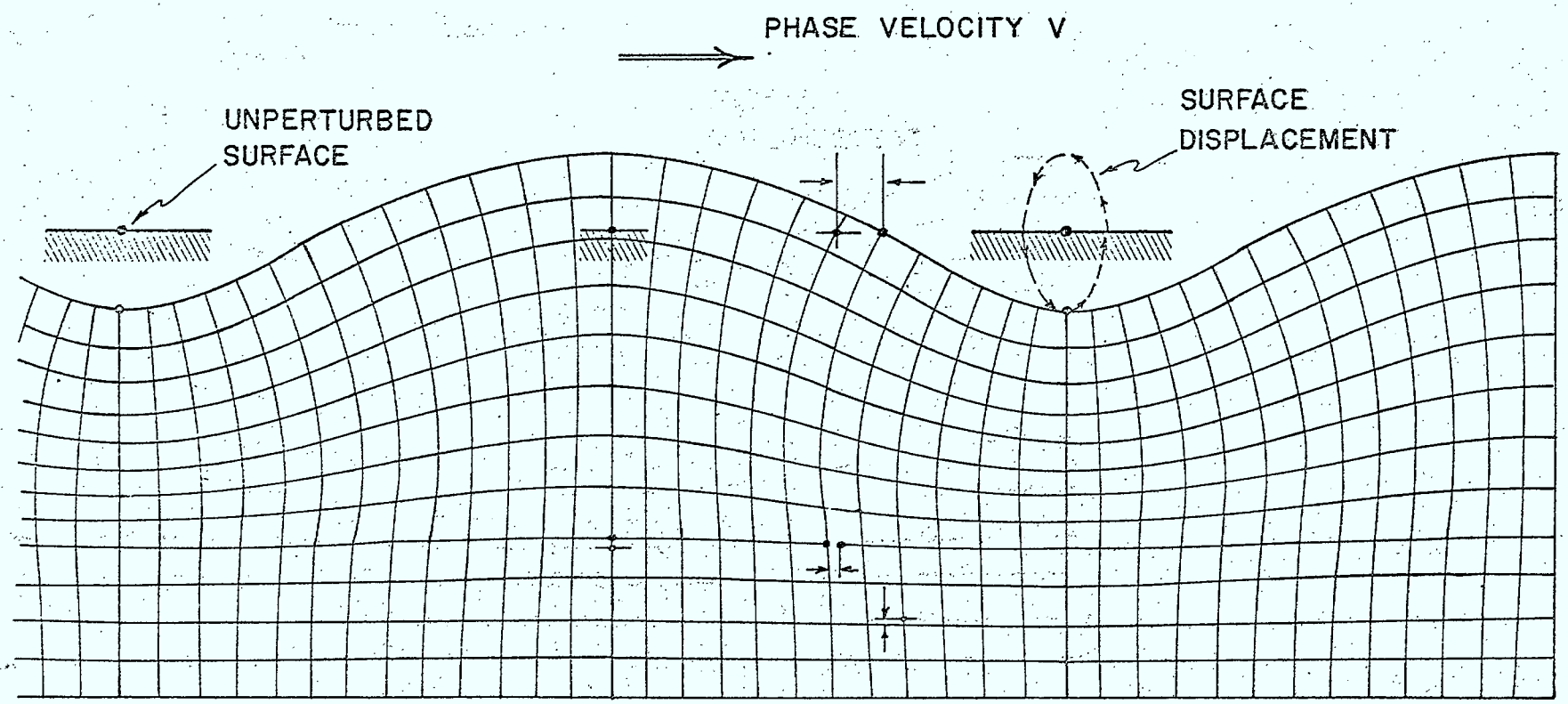


Fig. 2.4 Distortion of a square grid in sagittal plane due to surface-wave. Displacement amplitudes are exaggerated.

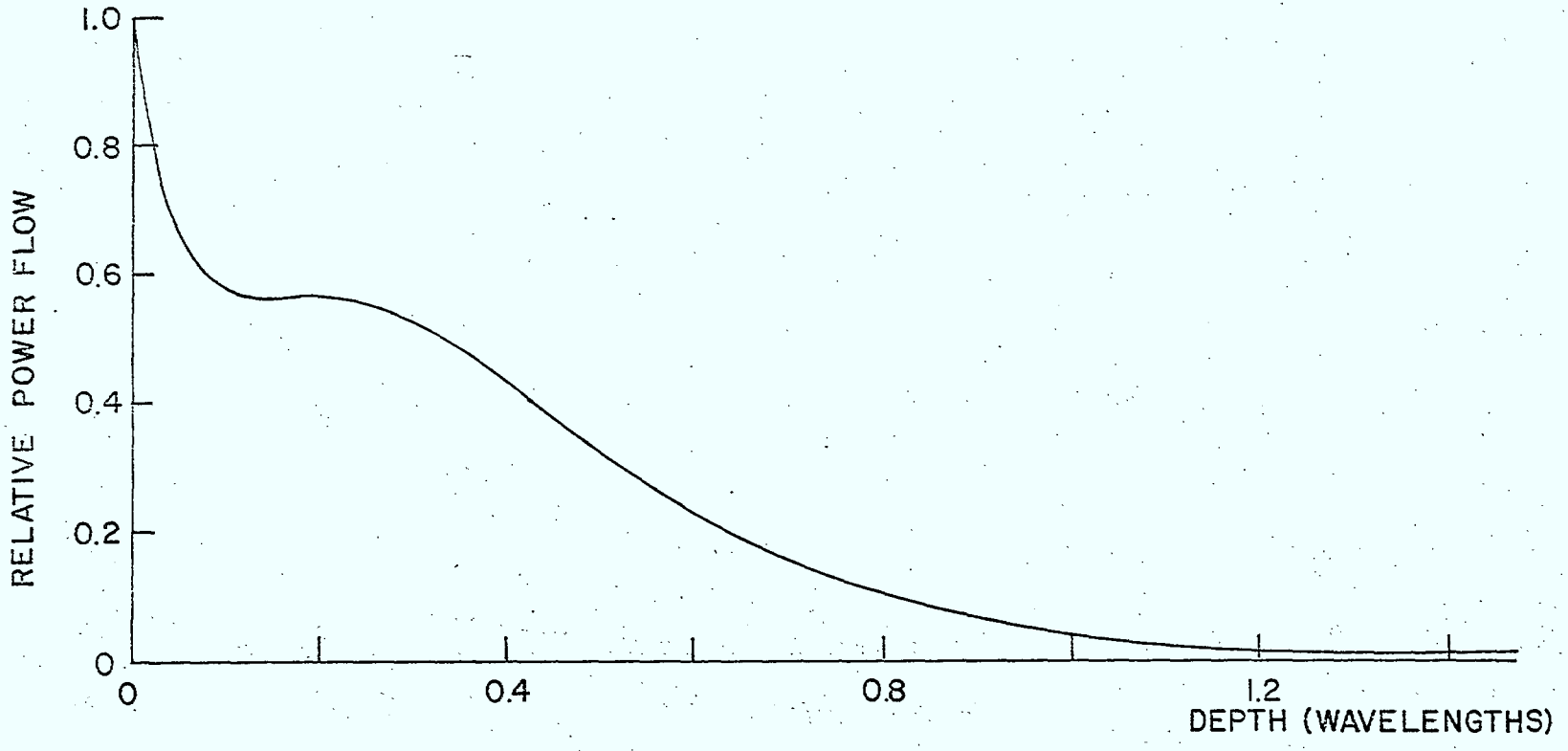


Fig. 2.5 Mechanical Poynting vector, power flow per unit area perpendicular to sagittal plane, as a function of depth below surface.

that single crystals, which are inherently anisotropic to some degree, must be used to minimize the attenuation. Attenuation is generally lower in a single crystal than in a polycrystalline sample of the same material, it increases as the square of the frequency and is much lower in some crystals than others, Chapter X. Anisotropy introduces many differences in detail to the propagation of surface waves (G 2) some of which are considered later but here it is noted only that with a single crystal substrate, the phase velocity varies with the crystal cut represented by the free surface and with the direction of the propagation vector on that surface. The direction of the propagation vector is fixed by the orientation of the source transducer and is almost invariably chosen along the extremum of the phase velocity versus angle curve on the chosen surface because for such a so-called pure-mode axis the energy transport is in the same direction as the propagation vector (93). Comparisons of various choices of crystal cuts and propagation directions are made and the effects of misalignment from a pure-mode axis discussed in Chapter X, but in all the remaining material on devices it is assumed that propagation is along a pure-mode axis.

2.2 Launching

While many techniques can be used for launching acoustic surface-waves, the only one used operationally to date has involved piezoelectric coupling with metal electrodes deposited on the surface (19). In a piezoelectric crystal, an electric field depending linearly on the amplitude of the amplitude of the strain is created in any region of distortion. It is obvious from Fig. 2.4 that the passage of a surface wave produces at any instant of time a complex but spatially periodic distribution of strain, the substrate is compressed in some directions in some regions, sheared in others, etc. Now if the substrate is a piezoelectric crystal this instantaneous pattern of displacement will produce a complex pattern of

electric field or electric potential, but a pattern which is also periodic in the direction of propagation. This periodic electric field pattern will propagate along with the mechanical disturbance. For a material of given elastic constants the presence of piezoelectricity reacts slightly onto the wave propagation characteristics and produces a slight increase of the surface-wave velocity. This velocity increase or piezoelectric stiffening is small but will be seen to be an important measure of the usefulness of the substrate for surface-wave devices (9).

The surface-wave on the piezoelectric substrate thus pulls along with it a periodic electric field distribution, the latter penetrates into the free space above the surface. If the wave passes under a pair of parallel finger-like electrodes plated onto the surface and perpendicular to the direction of the wave, a voltage periodic in time will be induced between the electrodes. The voltage amplitude will be a maximum if the electrodes are spaced one half wavelength apart. Thus we have a mechanism for detecting the surface-wave, and conversely, if an alternating voltage is applied across the electrodes from an external source we have a means of generating surface-waves. The prototype geometry for surface-wave devices was shown in Fig.2.1. Applying an alternating voltage to the input electrodes or interdigital transducer sets up a spatially periodic electric field which piezoelectrically generates the surface-wave and the latter propagates to a similar transducer which couples electrical energy out of the wave. The number of fingers in the transducers may vary from two to thousands depending on the application and similarly the length, spacing, shape, thickness and material of the fingers all affect the coupling and can be selected to suit the device characteristics required. The prototype interdigital transducers are considered quantitatively in the next chapter, while in later chapters more sophisticated designs for specific system needs are discussed.

III. TRANSDUCERS

The prototype electromechanical transducer for surface-waves consists of an interdigital comb structure of thin metal electrodes of the form shown in Fig.3.1a deposited onto the surface of a piezoelectric substrate (104). If an electrical voltage is applied between the two bus bars the instantaneous electric field sketched in Fig.3.1b will be established in the substrate, thus applying, by means of the piezoelectric coupling, a spatially periodic stress of period 2L to the surface region of the substrate. Varying the applied voltage periodically in time with an angular frequency ω, causes waves to radiate in both directions from each source region 2L long. When the frequency is $f_o = v/2L$ where v is the Rayleigh wave velocity in the direction of the transducer axis, then $2L = \lambda$, and the waves from each source region add in phase to create a surface-wave of width W propagating to the right which increases in amplitude linearly as it propagates from the first electrode on the left to the last electrode on the right, the resultant wave then leaves the transducer and continues as a free-surface Rayleigh wave. A similar wave is excited in the opposite direction. The behaviour at frequencies other than resonance $f_o = v/2L$ and the effects of varying W and L along the length of the transducer will be the theme of this and the subsequent chapters.

The electric field which is produced near the transducer fingers is illustrated in Figure 3.2 for the case of equal metallization to interelectrode space ($a = L/2$ in Fig.3.1a). The vertical and tangential components of field have different spatial distributions, both are large near the edges of the fingers and at relatively shallow depths

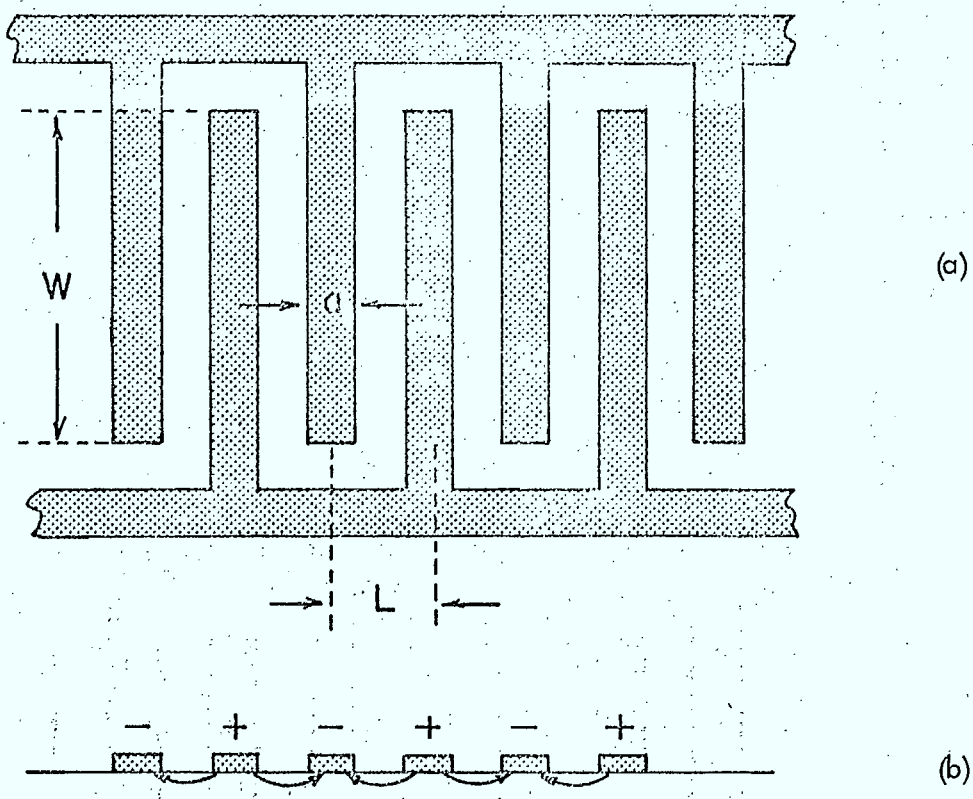


Fig. 3.1 (a) Geometry of an interdigital surface-wave transducer.

(b) Sketch of electric field in the substrate.

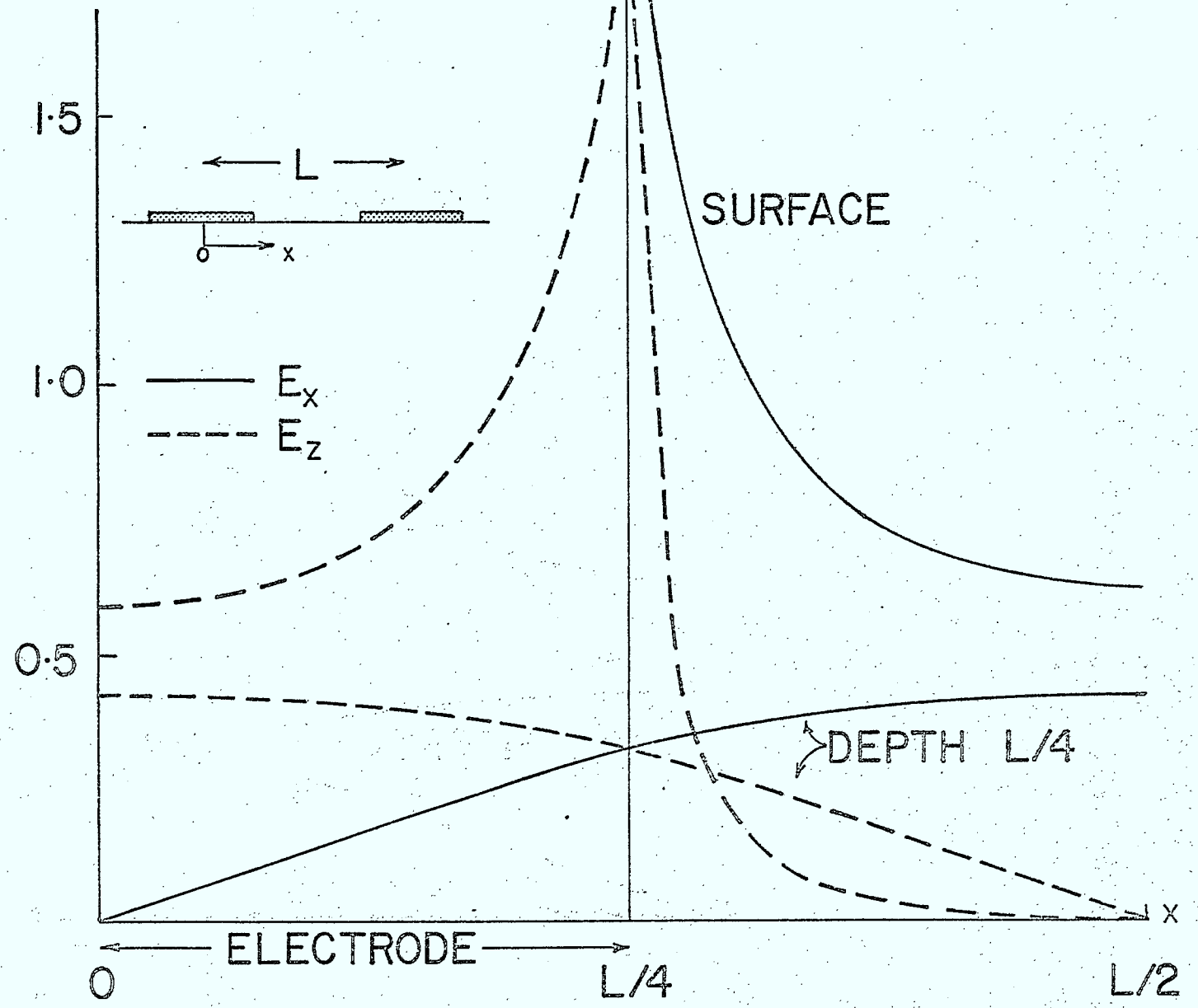


Fig. 3.2 Electric field components at the surface and L / 4 below the surface for one section of an interdigital transducer.

they become sinusoidal with space quadrature. These field components couple to a Rayleigh wave, itself having a complicated mechanical displacement distribution, through the various piezoelectric constants involved. It is thus not obvious what form of model should be used for the effective source distribution created by the interdigital transducer. Various approximations and models are available, several of which will be discussed here; which approximations and which model are to be employed in a given situation depend on the dominant specifications of the device being designed or analyzed and on the number of second-order effects which it is necessary to include. For example, in some cases it is advantageous to consider the source distribution as a set of spatial delta or impulse functions located at the finger edges, in others to consider the source distribution as sinusoidal with half-period L, and in yet others to assume that the source distribution is uniform over each successive distance L . The latter assumption forms the basis of a very useful equivalent circuit which will be discussed next.

3.1 Equivalent Circuit

If the transducer section of Fig.3.3a is modelled by the side-electroded bar of Fig.3.3b then the Mason equivalent circuit of Fig.3.3c applies (85, 86, 87). With a single-mode bar of large cross-section the characteristic impedance is $Z_o = \rho v A$ where ρ is the density and A the cross-sectional area, the acoustic to electric transformer has turns ratio $\varphi = W d/s$ with d and s appropriate piezoelectric and compliance constants (85). The circuit can be converted to a completely electrical one by introducing an electrical port equivalent to the acoustic port with

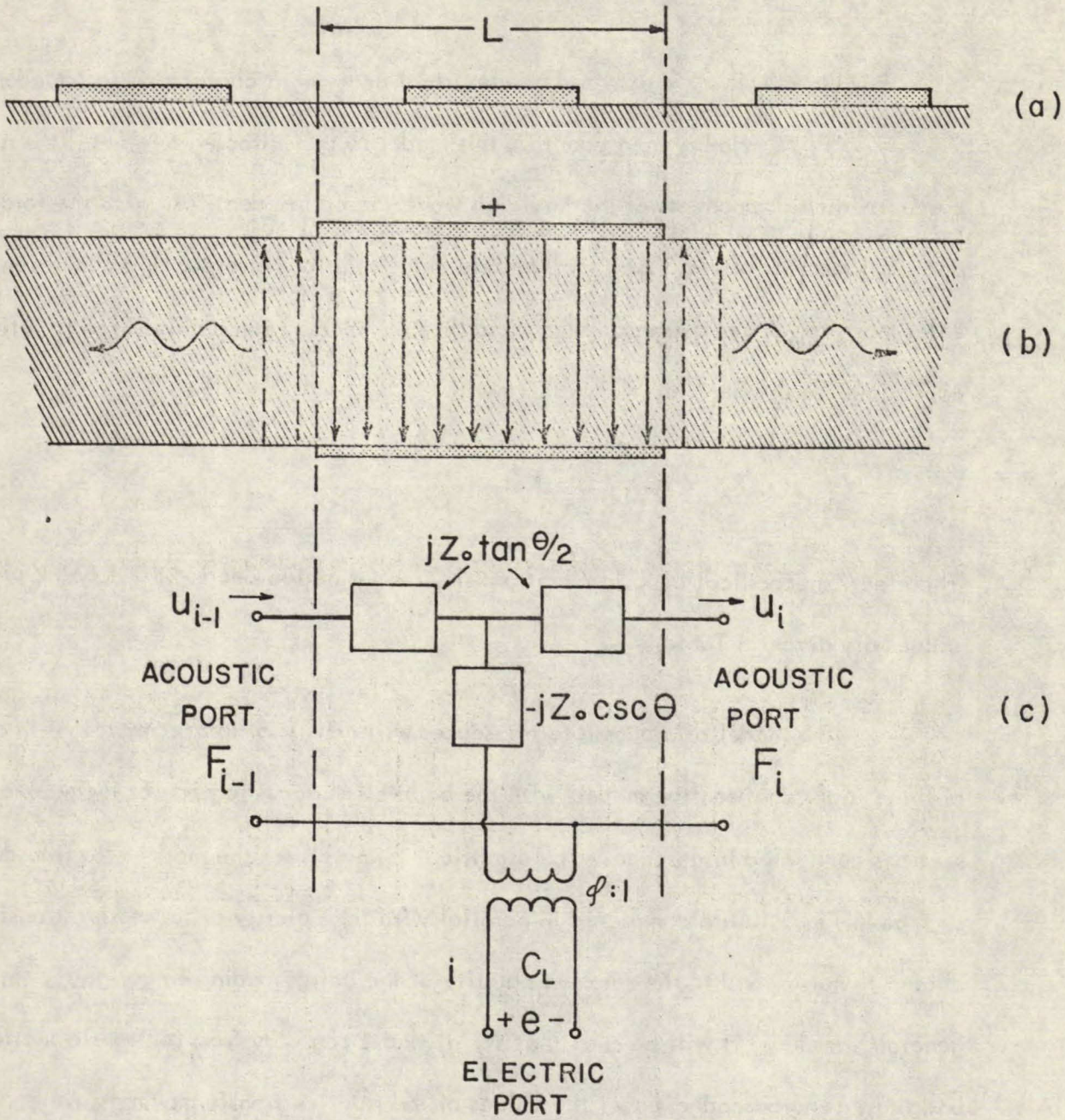


Fig. 3.3 (a) Cross-section of an interdigital transducer.

(b) Side-electroded bar model of one section of transducer.

(c) Mason equivalent circuit for length L of transducer based on (b). Z_o is mechanical impedance of bar, θ is transit angle $= \pi \omega / \omega_o = 2\pi L / \lambda$, ω_o is resonant frequency $= \pi v / L$, and C is static capacity of section.

$e_i' = F_i/\varphi$ and $i_i' = \varphi u_i$, and an electrical equivalent characteristic impedance $R_o = Z_o/\varphi^2$. However in adapting this model to the surface-wave case it is not evident which components of the Rayleigh wave should be identified with the force and displacement of Fig.3.3c. Fortunately only ratios of corresponding components enter most transducer design problems and thus the section can be modelled to a first approximation by the circuit of Fig.3.4 with

$$R_o = \pi/\omega_o C_L K^2 . \qquad (3.1)$$

The electromechanical coupling constant K^2 is defined in Section 10.1 and typical values are given in Table 3.1.

The model of a complete transducer will consist of interconnections of many of these individual section models with the equivalent acoustic ports of successive sections connected in cascade. The electrical terminals for the model of a transducer such as in Fig.3.1a are connected in parallel with the polarity of successive transformers chosen to correspond to the relative polarity of the corresponding finger pairs. In a general transducer it will be seen that W, a, and L can vary from section to section and thus in the cascaded model the values of the relative transformer turns ratio, static interelectrode capacity C, transit angle θ and characteristic resistance R_o appropriate to each section are used in the overall model.

The capacitance of one of the sections L can be calculated by a variety of techniques providing the adjacent sections are not too different from the one under consideration. If the metal fingers are long and very thin compared to their width,

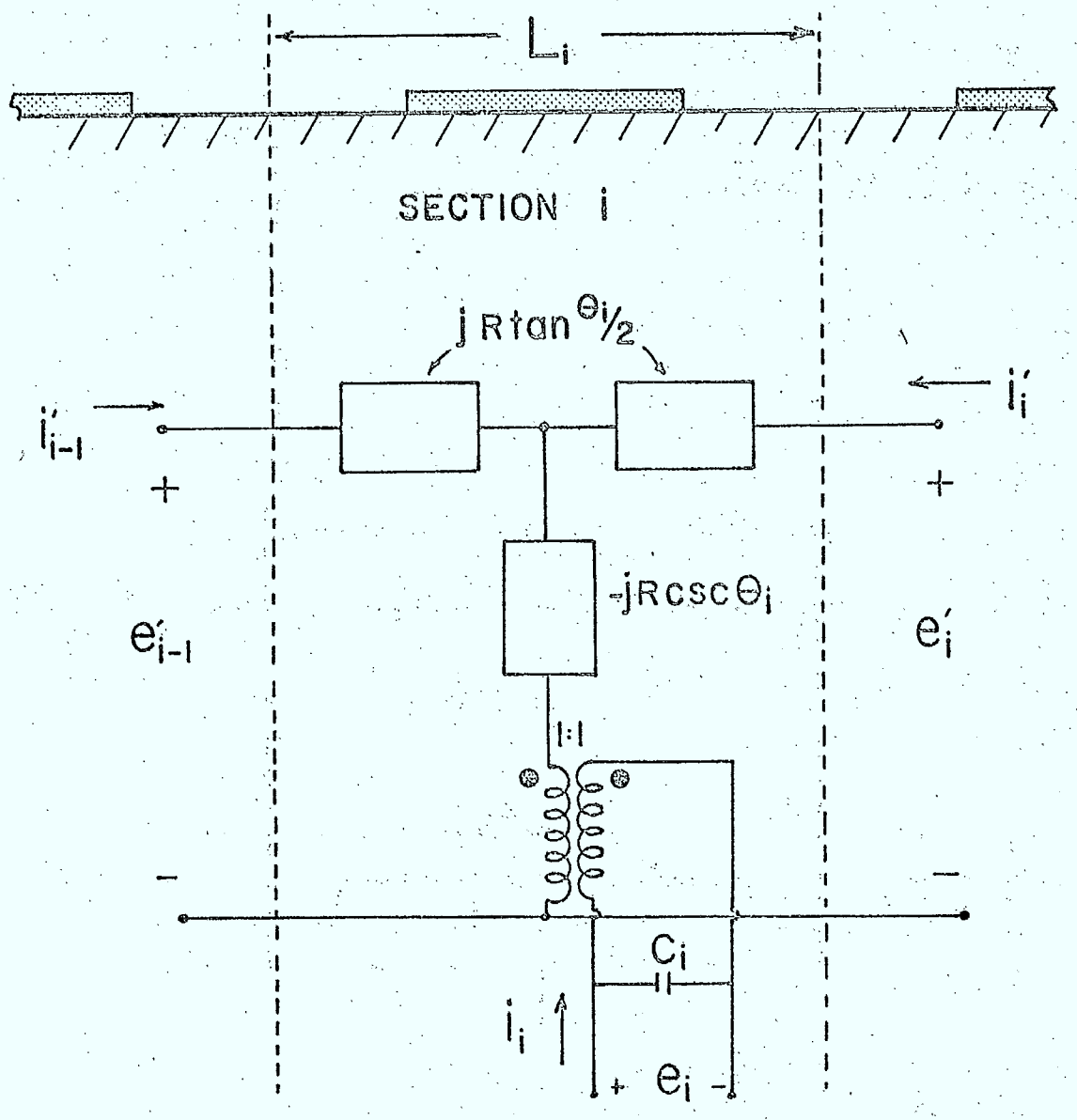


Fig. 3.4 Electrical equivalent circuit for ith section of transducer. R_o defined in Eq. 3.1, θ in Fig. 3.3.

the capacitance per unit length of a section L is given by the empirical relation (20)

$$C = (\epsilon_s + 1)\ \left[6.5\ (\frac{a}{L})^2 + 1.08\ (\frac{a}{L}) + 2.37\ \right] \qquad (3.2)$$

where "a" is the finger width, Fig.3.1a and ϵ_s is the effective dielectric constant for the chosen sagittal plane, for pure-mode cases ϵ_s is of the form $\epsilon_s = \sqrt{\epsilon_{11}\ \epsilon_{33} - \epsilon_{13}^2}$, where the ϵ_{ij} are dielectric constant tensor elements.

The admittance matrix description for the circuit of Fig 3.4 is (85)

$$\begin{bmatrix} i'_{i-1} \\ i'_i \\ i_i \end{bmatrix} = j G_o \begin{bmatrix} -\cot\theta & \csc\theta & -\tan\theta/2 \\ \csc\theta & -\cot\theta & -\tan\theta/2 \\ -\tan\theta/2 & -\tan\theta/2 & 2\tan\frac{\theta}{2} + \omega C/G_o \end{bmatrix} \begin{bmatrix} e'_{i-1} \\ e'_i \\ e_i \end{bmatrix} \qquad (3.3)$$

where $G_o = R_o^{-1}$.

3.2 Uniform Transducer

For the uniform transducer there are N identical sections each of the same length 2L and width W connected as in Fig.3.1a. Thus the equivalent circuit consists of 2N sections of the type shown in Fig.3.4 with the acoustic ports in cascade and the electrical ports in parallel with the transformers of alternate

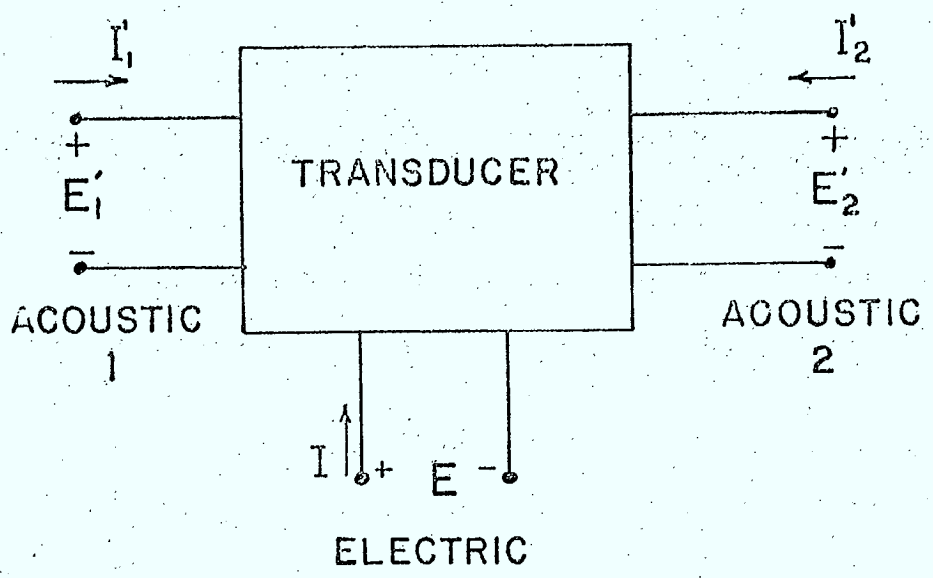
sections having opposite polarities. The resultant admittance matrix for the N - pair uniform transducer of Fig.3.5 becomes

$$\begin{bmatrix} I'_1 \\ I'_2 \\ I \end{bmatrix} = jG_o \begin{bmatrix} -\cot 2N\theta & \csc 2N\theta & -\tan \theta/2 \\ \csc 2N\theta & -\cot 2N\theta & -\tan \theta/2 \\ -\tan \theta/2 & -\tan \theta/2 & 4N \tan \frac{\theta}{2} + \omega C_T/G_o \end{bmatrix} \begin{bmatrix} E'_1 \\ E'_2 \\ E \end{bmatrix} \quad (3.4)$$

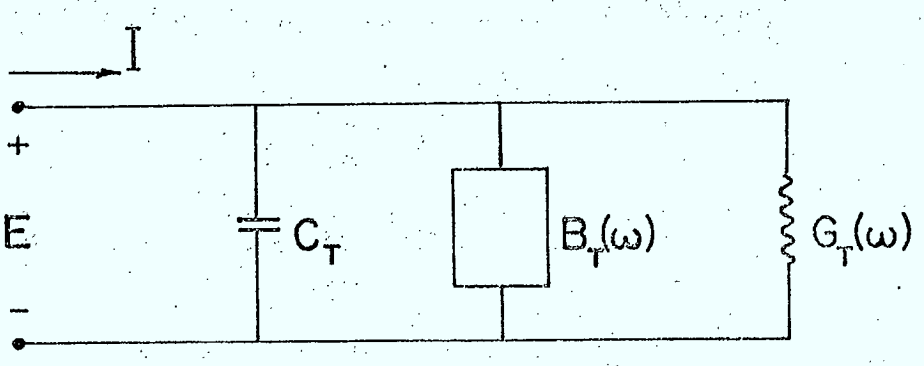
where $C_T = 2NC$ and $\theta = 2\pi L/\lambda = \pi\omega/\omega_o$

If there are no reflections from discontinuities in the acoustic path to the right or left of the transducer, the two acoustic ports are terminated in their characteristic impedance R_o so that the equivalent circuit as viewed from the electrical port becomes that of Fig.3.5b with

$$\begin{aligned} G_T(\omega) &= 2G_o \tan^2 \frac{\theta}{2} \sin^2 N\theta \\ B_T(\omega) &= G_o \tan \frac{\theta}{2} \left[4N + \tan \frac{\theta}{2} \sin 2N\theta \right] \end{aligned} \quad (3.5)$$



(a)



(b)

Fig. 3.5 (a) Three-port model of complete transducer.

(b) Equivalent circuit as seen from electrical terminals.

For frequencies near the fundamental resonance $\omega = \omega_o$ with $\omega_o = \pi\ v / L$ the radiation conductance is

$$G_T(\omega) = 8\ N^2\ G_o\ \left(\frac{\sin X}{X}\right)^2$$

$$\text{and} \qquad B_T(\omega) = 8\ N^2\ G_o\ \frac{\sin 2X - 2X}{2X^2} \tag{3.6}$$

$$\text{where} \qquad X = N\pi\ \frac{(\omega - \omega_o)}{\omega_o}$$

With the equivalent circuit of Fig.3.5b the power dissipated in the radiation conductance G_T represented the acoustic power radiated as a surface-wave from the transducer, one half of the power leaving from each end of the transducer. Figure 3.6 shows the radiation conductance and susceptance as a function of the normalized differential frequency. Note that the radiation susceptance here is zero at resonance and that to obtain the total input susceptance, ωC_T must be added.

Thus as would be expected from the truncated, periodically alternating nature of the source distribution represented by the uniform transducer, the power radiated is of band-pass form and the bandwidth is inversely proportional to the length of the transducer.

While shaping of the transducer band-pass characteristics will be discussed in detail later, it can be noted here that if the only matching used is a series inductor to resonate the electrode capacity C_T at ω_o, the maximum overall bandwidth is obtained when the number of finger-pairs is given approximately by (86)

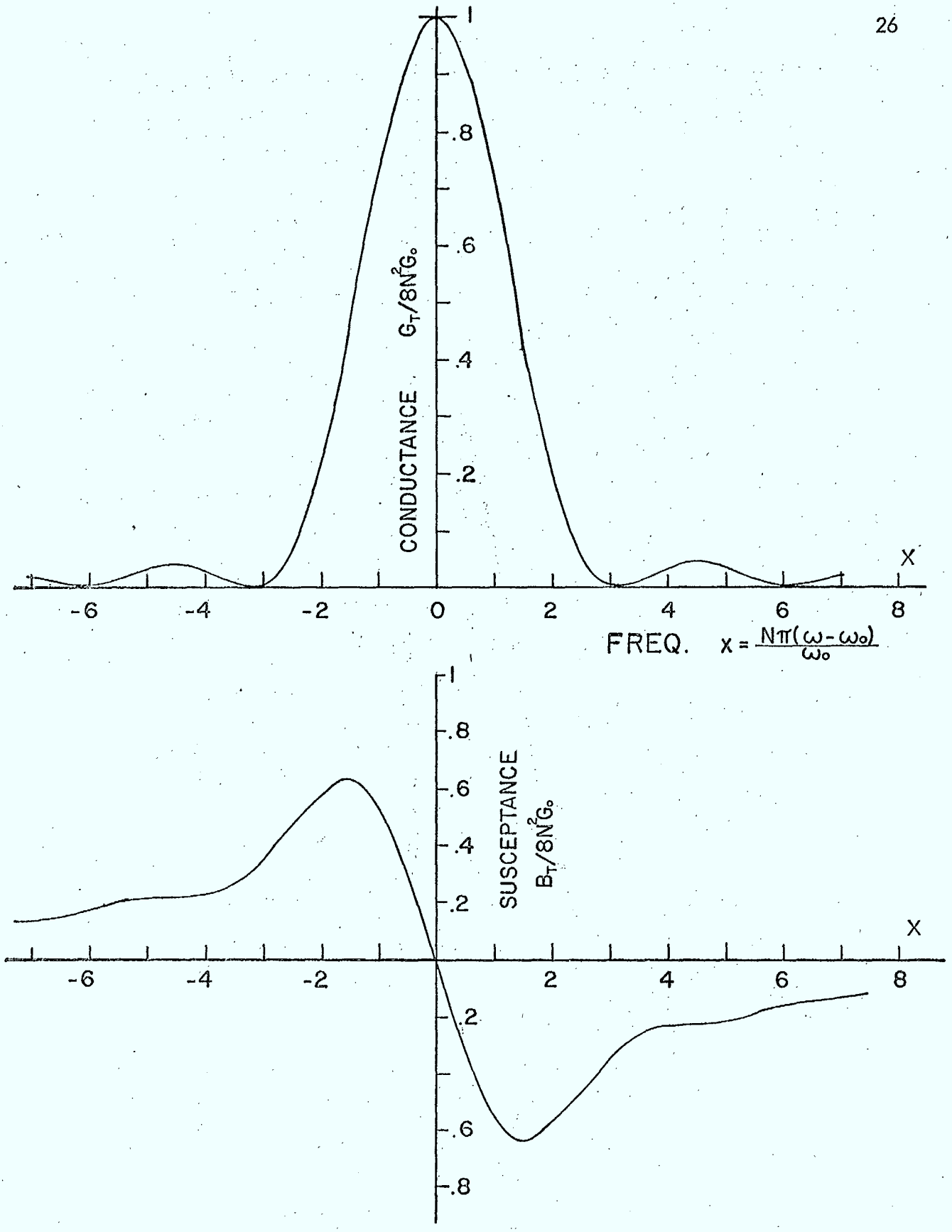


Fig. 3.6 Input admittance of a uniform interdigital transducer, Eq. 3.6, as a function of normalized frequency.

TABLE 3.1

OPTIMUM TRANSDUCER DESIGN FOR VARIOUS SUBSTRATES

Piezo-electric	Cut and Propagation Direction	$K^2/2$ %	N_{opt}	W/λ_o	$\frac{\Delta\omega}{\omega_o}$
$LiNbO_3$	YZ	2.46	4	108	0.24
$Bi_{12}GeO_{20}$	[110], [$\bar{1}$10]	1.15	6	183	0.17
ZnO	XZ	0.56	8	99	0.12
Quartz	YX	0.11	19	53	0.053
PZT		2.15	4	-	0.23
CdS	XZ	0.31	12	54	0.09
$LiTaO_3$	ZY	0.82	7	31	0.14

$$N_{opt}^2 = \frac{\pi}{4\,K^2} \tag{3.7}$$

Having chosen the number of finger pairs, the length of the fingers W can be chosen to match the source resistance at resonance. Table 3.1 (86) shows the optimum number of finger pairs and the transducer width in wavelengths for a 50-ohm source for several different substrates; also shown is the half-power fractional bandwidth which is approximately $1/N$.

The preceding description considers the principal performance characteristics of the prototype uniform transducer. Many variations of transducer geometry to produce other than $(\sin x)/x$ frequency responses will be analyzed in later sections and in the course of these discussions several second-order effects will be considered. The latter include reflections, diffraction losses, anisotropic effects, harmonic and bulk-wave response etc. It is advantageous to include at this point an introduction to two types of reflection, reflection from the transducer as a whole producing echoes on the acoustic path, and reflection internal to the transducer by the characteristic impedance change between electroded and unmetallized regions.

If an acoustic wave is incident from the left onto the transducer represented by Fig.3.5a and there is no reflection from the acoustic load to the right, then the acoustic ports are respectively fed from and terminated in the characteristic impedances R_o. In general the power reflected from one input port and the power delivered to the other two ports can be calculated from the overall admittance matrix of Eq.3.4 for any choice of terminations and expressed as a power scattering matrix with coefficients (87)

$$p_{ij} = \frac{P_i}{P_{av\,j}} \tag{3.8}$$

where P_i is the power transmitted or reflected from port i and $P_{av\,j}$ is power available from a matched generator at port j. By reciprocity $p_{ij} = p_{ji}$ and for the symmetrical case of a uniform transducer $p_{22} = p_{11}$ and $p_{23} = p_{13}$.

Thus in our example of an acoustic wave incident on port 1, if the electrical port is loaded by an admittance which resonates the electrode capacity at the synchronism frequency ω_o, i.e. $Y_L = G_L - j\,\omega_o\,C_T$ then at ω_o

$$p_{11} = \frac{1}{(1+b)^2} \qquad p_{12} = \frac{b^2}{(1+b)^2} \qquad p_{13} = \frac{2\,b}{(1+b)^2} \tag{3.9}$$

where $b \equiv G_L / 8\,N^2\,G_o$. The maximum conversion to electrical power then occurs for

$$G_L = 8\,N^2\,G_o \tag{3.10}$$

and one half of the incident available power is absorbed in the electrical load, one quarter is reflected at the acoustic input port and one quarter is radiated from the other acoustic port.

It is interesting to note also from Eq.3.9 that for G_L infinite (shorted electric port) all of the incident power is transmitted through to the output acoustic port and that for G_L zero, $Y_L = -\omega_o\,C_T$, and all of the incident power is reflected at $\omega = \omega_o$.

3.3 Finger Reflections

Plating the surface of a piezoelectric changes the velocity of the Rayleigh wave in the plated region, in part due to the electric shorting of the surface and in part due to the mechanical loading. Thus in one section of the transducer of Fig.3.1a the characteristic impedance for the part under the finger is different from

that in the space between. The mechanical discontinuity can be minimized by using a plating that is as thin as allowed by the requirement of electrical continuity in the fingers and by using a light metal such as aluminum. Note though that decreasing the electrode thickness increases the resistance of the fingers and hence the ohmic losses in the transducer. The electrical discontiniuty is more inherent because the ratio of the characteristic impedance R_o in the metallized region to that of the unmetallized R'_o is approximately

$$\frac{R'_o}{R_o} = \frac{1}{1 + K^2/2} \tag{3.11}$$

where K^2 is an electromechanical coupling factor characteristic of the piezoelectric (Chapter X) whose values will be seen to be 0.1 to 5% for the piezoelectrics of interest. This difference in impedances can be important for high coupling materials such as the YZ - Li $Nb O_3$ which is a very common surface-wave substrate.

The electric shorting of the finger regions has two effects on the frequency response of the transducer, the acoustic mismatches at the finger edges produce reflections which distort the shape of the transfer function or the driving point electrical impedance, and the shift of the "average" velocity of propagation in the transducer region which results in a slight shift of the centre frequency of the transducer. In most applications the reflections and the distortion of the band-pass characteristic are the more undesirable of the two effects.

In order to reduce third harmonic excitation which is zero for $a/L = 2$ and to relax fabrication tolerances as much as possible, transducers of the form in

Fig.3.1 are usually fabricated with electrode-to-space ratios of the order of unity. With such ratios the acoustic reflections from all of the successive discontinuities add in phase at resonance as illustrated in Fig.3.7 (36), with the result that even small discontinuities produce large cumulative reflection effects. The acoustic discontinuities can be included in the model of a finger of the interdigital transducer by breaking the section of Fig.3.4 into three parts corresponding to the metallized half-fingers and the intervening unmetallized section to give the model of Fig.3.8. The 3-port admittance matrix of this circuit can be determined explicitly (87).

If the reflections caused by the addition of the individual reflections of Fig.3.7 are important and if narrower fingers can be fabricated in the application under consideration, the reflections can be caused to cancel by splitting each finger into two as illustrated in Fig.3.9 so that the total reflection from each pair of fingers is zero, Fig.3.9b (G6, p.472),(G5, p.343). The same model shown in Fig.3.8 can be used for each finger of the split-finger transducer of Fig.3.9 but in performing the electrical connections the polarity of the transformers must be chosen to correspond to the appropriate bus connections of the fingers. Note that there is no field in the region between the two parts of a split finger and thus the gross behaviour of the transducer is that of one composed of fingers of width a, Fig.3.9a, and relatively narrow gaps.

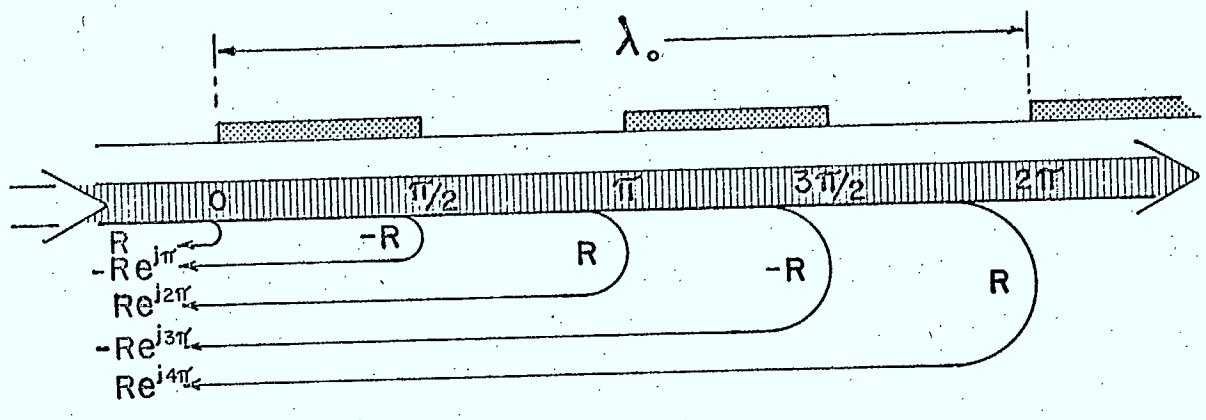


Fig. 3.7 First-order reflections from the edges of the transducer fingers.

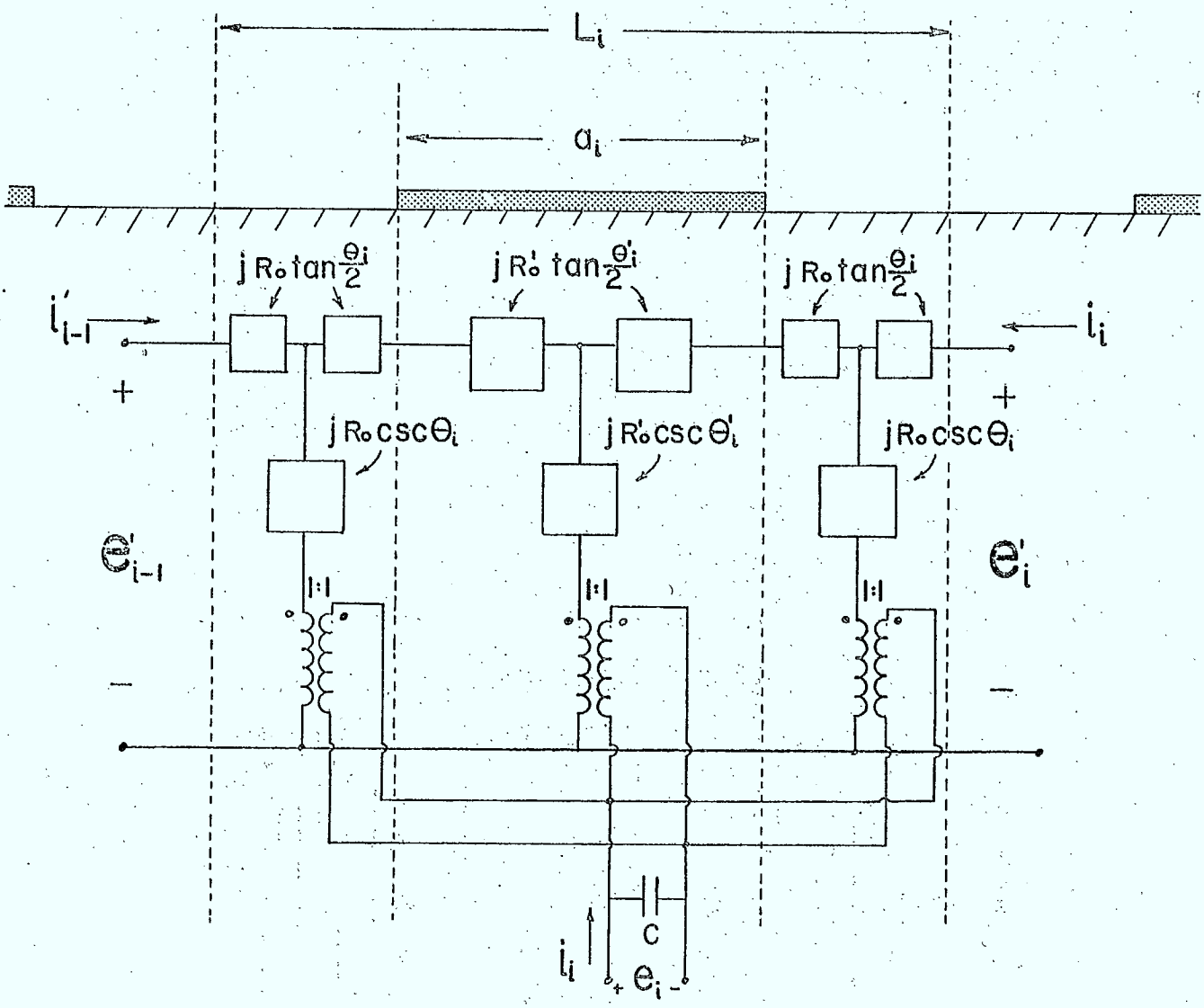


Fig. 3.8 Equivalent circuit of one section of a transducer including mismatch at finger edges. $\theta_i' = 2\pi a_i / \lambda'$, $\theta = \pi (L_i - a_i) / \lambda$, $\lambda = v/f$, $\lambda' = v'/f$, $v' = v/(1 + K^2/2)$ and the resonant frequency

$$f_i = \frac{1}{2}\left[\frac{a_i}{v} + \frac{L_i - a_i}{v'}\right]^{-1}.$$

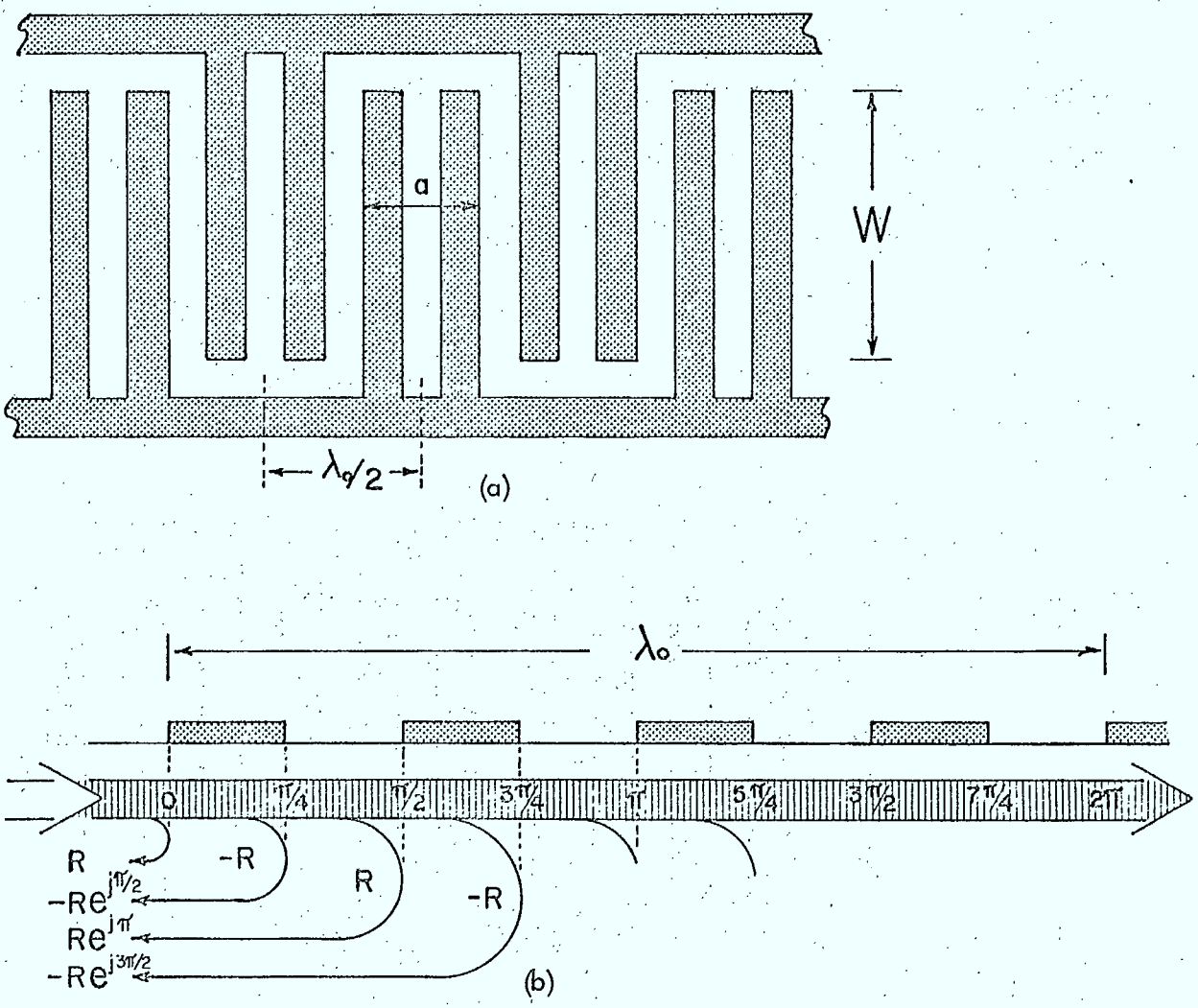


Fig. 3.9 (a) Geometry of a split-electrode transducer.

(b) Reflections from finger edges of a split-electrode transducer.

3.4 Unidirectional Transducers

The transducers discussed so far have been bidirectional so that when an electrical voltage is applied acoustic waves are radiated from each end; for symmetrical interdigital transducers these two waves are identical. Thus when two bidirectional transducers are used as an input-output pair there is a basic 6 db insertion loss due to the bidirectionality alone. In many applications, especially where the signal processing is being done at intermediate frequencies, this insertion loss is undesirable but tolerable. However in applications such as receiver front ends or large dynamic range systems this loss is not tolerable; and in a far wider range of applications the bidirectionality is detrimental to performance to a far larger extent than the 6 db would imply.

It was noted in Eq.3.9 that if a simple transducer were matched to produce maximum electrical power from a given incident acoustic wave then one quarter of this incident power becomes a reflected acoustic wave, and when re-reflected from a matched input transducer gives rise to "triple transit echoes". The only way to perfectly match the acoustic port of a bidirectional transducer is to reduce the electrical output to zero. Thus there is a double incentive to search for unidirectional transducers, the insertion loss is reduced by 6 db and for a given insertion loss the acoustic reflections can be smaller, that is the unidirectional transducer can be matched at the acoustic port.

For narrow-band devices it is relatively simple to produce unidirectional transducers. For example, a second, passive transducer with its static capacitance resonated out by an inductance at a frequency ω_o will completely reflect (Eq.3.9) all the acoustic energy incident on it at frequency ω_o from the driven transducer, and if the centre lines of the two transducers are separated by an odd multiple of $\lambda_o/4$ the pair of arrays constitute a unidirectional transducer (86). But because the path length between centre lines is necessarily many wavelengths, the front-to-back ratio of this composite unidirectional transducer is very frequency dependent.

One broadband method of producing unidirectional transducers uses the multistrip coupler and is discussed in Chapter IX. Another method is to provide more than the two half-wavelength spaced samples per wavelength of a standard interdigital transducer. In which case the electric field pattern produced by the driving voltage can be made effectively to progress along the length of the transducers in a manner analogous to the translating magnetic fields in a linear induction motor. For example if the three-fingers-per-wavelength array of Fig.3.10a (27) were driven by a three-phase voltage at the three electrical input terminals, the electric field pattern would progress to the right or to the left depending on the ordering of the three phases and would not be the standing field pattern of the normal transducer. Here the bandwidth of the transducer is comparable to that of a normal transducer of the same length. The increased number of fingers per wavelength increases the fabrication tolerances somewhat, but the more difficult

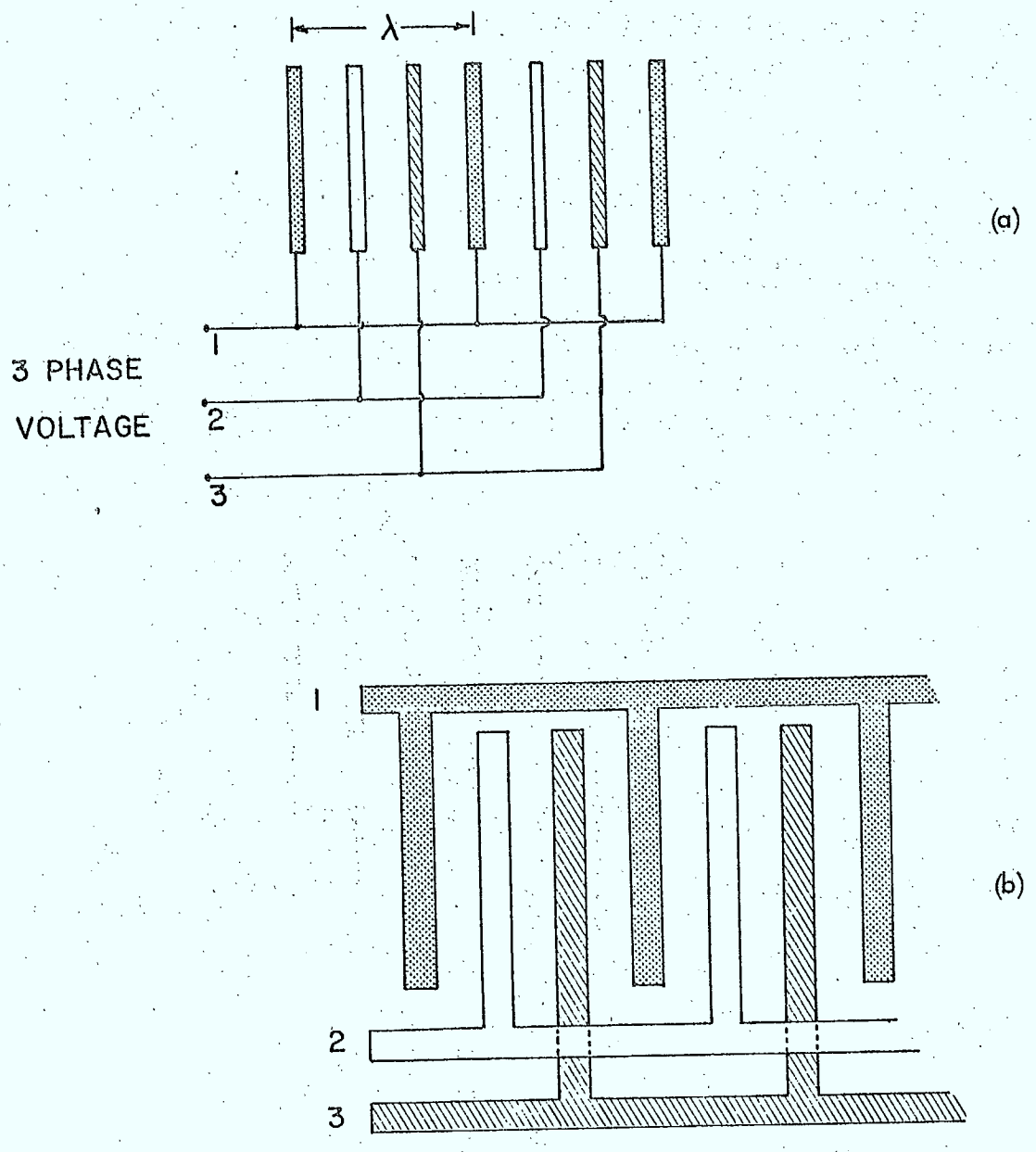


Fig. 3.10 (a) Schematic operation of a three-phase unidirectional transducer.

(b) Geometry of a three-phase unidirectional transducer.

fabrication problem is introduced by the necessary crossovers required by the third input lead as shown in Fig. 3.10b. The other serious difficulty with this configuration lies in the provision of a matching network to produce the balanced three phase signal over a wide range of frequencies. Transducers of this form giving more than 20 db directivity over a 20% bandwidth have been realized (27).

3.5 Harmonic Operation

It is apparent from Fig. 3.2 which gives the field distribution near the electrodes of an interdigital transducer that there is a large spatial harmonic content to this field (19, 99). Thus the coupling from the input electric source to the surface-wave should be large at frequencies harmonically related to the fundamental. By symmetry such coupling will exist only for the odd harmonics, and analysis of the field shows that the relative amplitudes of the harmonic components for a given amplitude of driving voltage are as shown in Fig. 3.11. It is seen that for a metallization ratio a/L of 0.5, that is equal width strips and spaces, the fifth harmonic is the lowest one excited and the coupling for it is about one fifth the value for the fundamental. Extremely large or extremely small values of metallization ratio give third harmonic amplitudes of the order of one half that available for the fundamental.

Harmonic operation of a transducer has the advantage of allowing much larger finger widths and thus reduces the fabrication tolerances. It should be

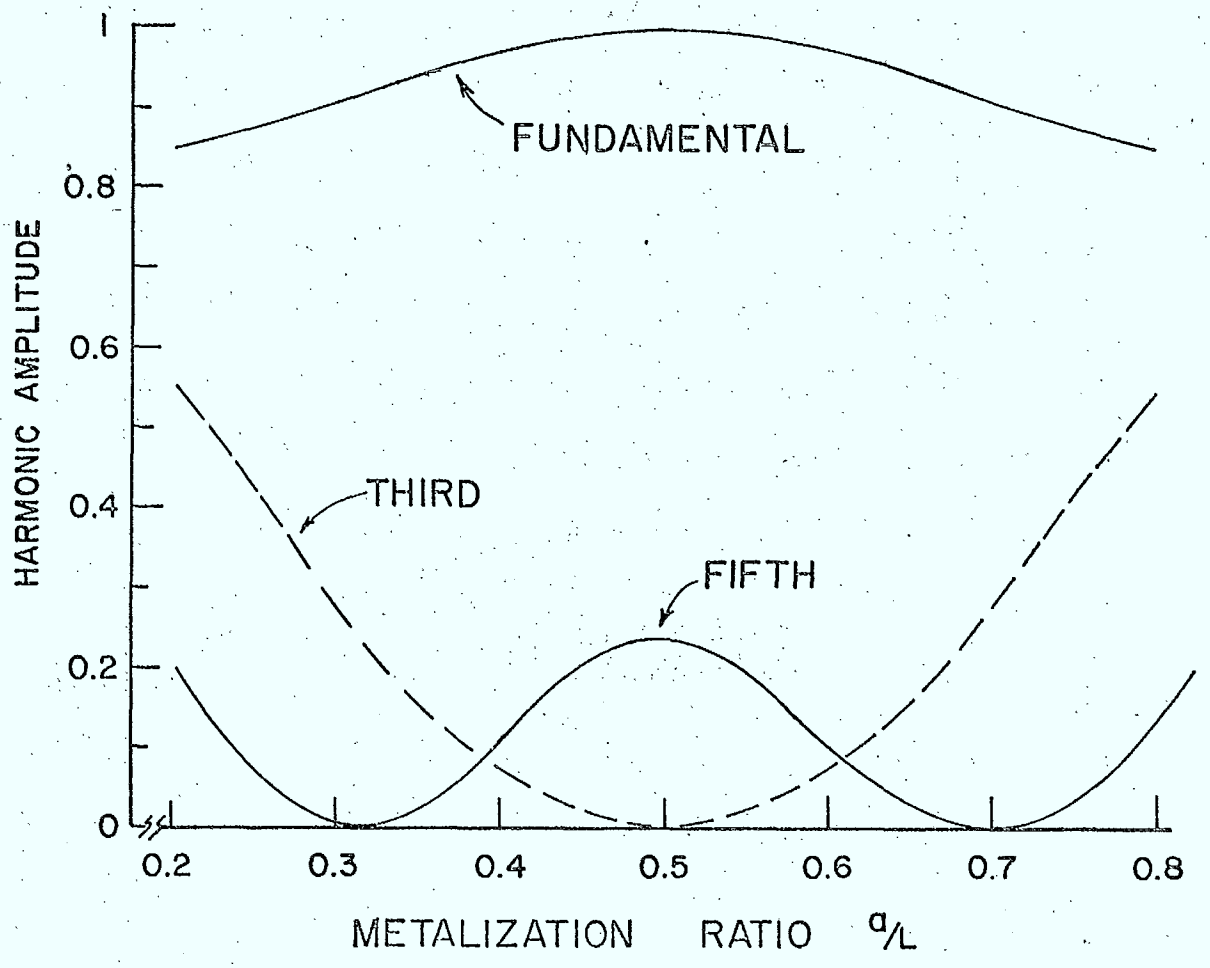


Fig. 3.11 Relative amplitude of excitation of fundamental and harmonics of a uniform transducer as a function of metallization ratio, normalized to fundamental value at $a / L = 0.5$ (4).

noted however that this relaxed tolerance is really only valid for the fifth harmonic operation not the third because in the latter case the extreme metallization ratios reimpose tolerances comparable with fundamental operation. With fifth harmonic operation the coupling is low.

It is questionable whether harmonic operation will see extensive use in surface-wave devices in the near future because fundamental devices can be fabricated up to frequencies very close to where propagation losses provide the frequency limitation.

3.6 Finger Resistance

In the previous sections it has been implied that the metal fingers of the transducer are perfectly conducting. However the film thickness is limited by the need to avoid mechanical discontinuities on the substrate surface. Thus in fact the fingers have a finite distributed resistance, and thus the finger and its interelectrode gaps form along its length more of a distributed RC line than the lumped capacity in the model of Fig. 3.3. The effect of this distributed resistance is to reduce the acoustic power radiated for a given total input electric power. If the ratio of these powers is called the transducer efficiency, then this efficiency will drop with increasing sheet resistivity of the metal or increasing finger length in the manner sketched in Fig. 3.12. It is seen that for wide aperture

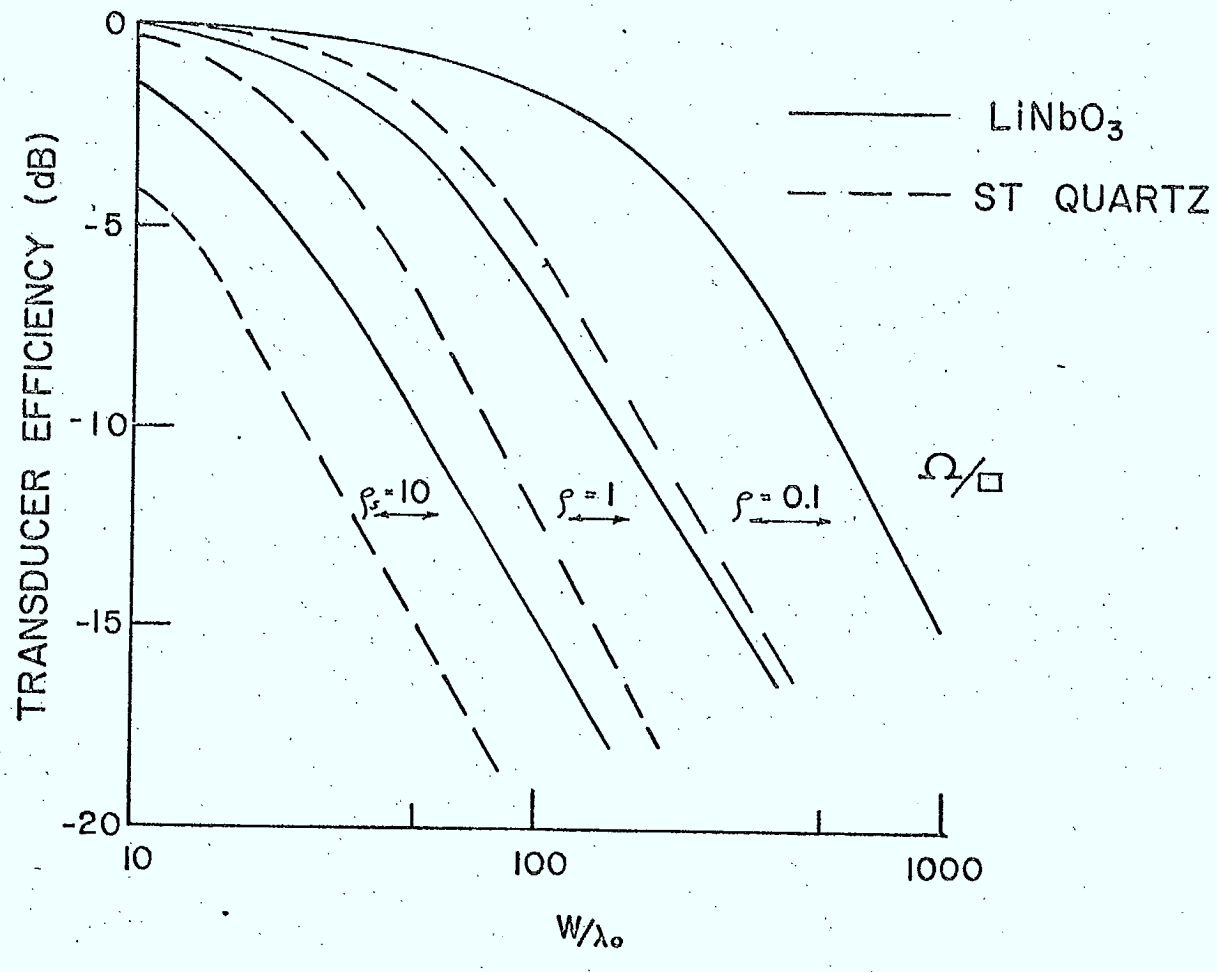


Fig. 3.12 Transducer efficiency in decibels as a function of aperture for different sheet resistivities (46).

transducers the resistive losses can become important for sheet resistivities larger than some tenths of an ohm per square. This limitation must be kept in mind in face of the desire to keep the films as thin as possible, because it is difficult to maintain bulk-like conductivities in very thin films.

IV. BANDPASS FILTERS

It was seen in Section 3.1 that the region near one finger of an interdigital transducer could be modelled as a three-port network, Fig.3.4, in which the parameters were expressed in terms of the various transducer dimensions and the properties of the substrate. In Section 3.2, the terminal properties of a uniform transducer of N - finger pairs, Eq.3.4, were determined by connecting 2 N of the identical individual sections in cascade. This amounts to the connecting of alternate fingers to the two bus bars of the transducer. The amplitude of the wave launched from a uniform interdigital transducer near the fundamental resonance is proportional to the frequency term $(\sin X) / X$ where $X = N \pi (\omega - \omega_o) / \omega_o$ and $\omega_o = \pi v / L$, Eq.3.6, assuming the input reactance is tuned out, and the beam has a width equal to the overlap length of the fingers of the comb. If a second matched uniform transducer of the same finger length is used to receive the wave launched by the first, the overall transfer function is the product of the two individual responses and thus the combination of the two transducers gives a band-pass transfer characteristic, the details of this characteristic depending on the values of N and L of the two uniform combs.

However it is evident that there are several internal parameters which can be varied in either or both of the transducers. For example the resonant frequency ω_o can be made different for different regions of the comb by varying L from finger-pair to finger-pair; similarly the ratio of finger width to interelectrode space can be varied changing the capacitance and the characteristic impedance, Eq.3.1;

and the finger overlap length can be varied from pair to pair varying the width of the acoustic beam radiated from the different portions of the transducer with their different resonant frequencies; even the relative phase of the signal radiated from different parts of a transducer can be reversed by interchanging the order in which finger pairs are connected to the two bus bars. Thus it is obvious that a vast number of different frequency response characteristics can be realized using the parameters available in a pair of non-uniform interdigital transducers.

The frequency response of any specified transducer can be determined by appropriate interconnection of the model sections of Fig.3.4 or 3.8 appropriately modified section by section to suit the parameters as they change from finger-pair to finger-pair. There exists a computer program which will calculate most of the characteristics of a transducer of quite a range of useful geometries (97). It provides for example the insertion loss, group delay and driving point impedances of pairs of transducers as a function of frequency. The program also calculates the effects of electrical matching networks and the departure of the phase vs. frequency curve from a linear or quadratic dependence. The impulse response with the matching network and the amplitude of a generated or reflected acoustic wave are also available. The most conveniently specified geometries are those of the forms to be discussed in this chapter and the succeeding two.

Thus while the analysis of non-uniform transducers is complicated in detail, it is relatively explicit and straight-forward. However in systems applications the designer tends to be faced initially not with the analysis problem but with the much

more difficult problem of synthesizing a component to meet a given set of specifications and tolerances. A unique "best" method for the synthesis of surface-wave transducers has not yet become apparent and different approaches are used for different applications, and even for a given application and specification there are often competing design procedures. Some general concepts regarding synthesis for band-pass filters will be considered in the next section while dispersive filters will be treated in Chapter V and in particular, non-dispersive combinations of dispersive elements are outlined in Section 5.3.

4.1 Synthesis Concepts

It is important to realize first that the surface-wave transducer is fundamentally a transversal filter as illustrated in Fig. 4.1 (63, 90). In Fig. 4.1a a straight-crested surface-wave is incident onto the output transducer which has pairs of electrodes spaced at varying distances along the length of the transducer. The voltage induced in the load by a given finger pair, that is the transformer turns ratio in Fig. 3.4, will be proportional to the overlap length. Since the energy tapped from the propagating wave at each finger is small and independent of frequency over a wide bandwidth, the output transducer is analogous to the general transversal filter of Fig. 4.1b. The delays $\tau_1, \tau_2 \ldots \tau_N$ correspond to the dispersionless propagation delay between taps and the weighting factors $W_1, \ldots W_N$ correspond to the overlap length of the fingers of the tap (see Eq. 4.6 below) with the sign of the weighting factor depending on the bus connections of the finger pair. If the nth tap is located at x_n, the total delay to it is x_n/v, and the frequency response of the transducer is proportional to (95), (G 4, p.115)

$$H(\omega) = \sum_{n=1}^{N} W_n e^{j\omega x_n/v} \quad . \tag{4.1}$$

Varying the overlap length for the fingers along a transducer is referred to as "apodizing" (96).

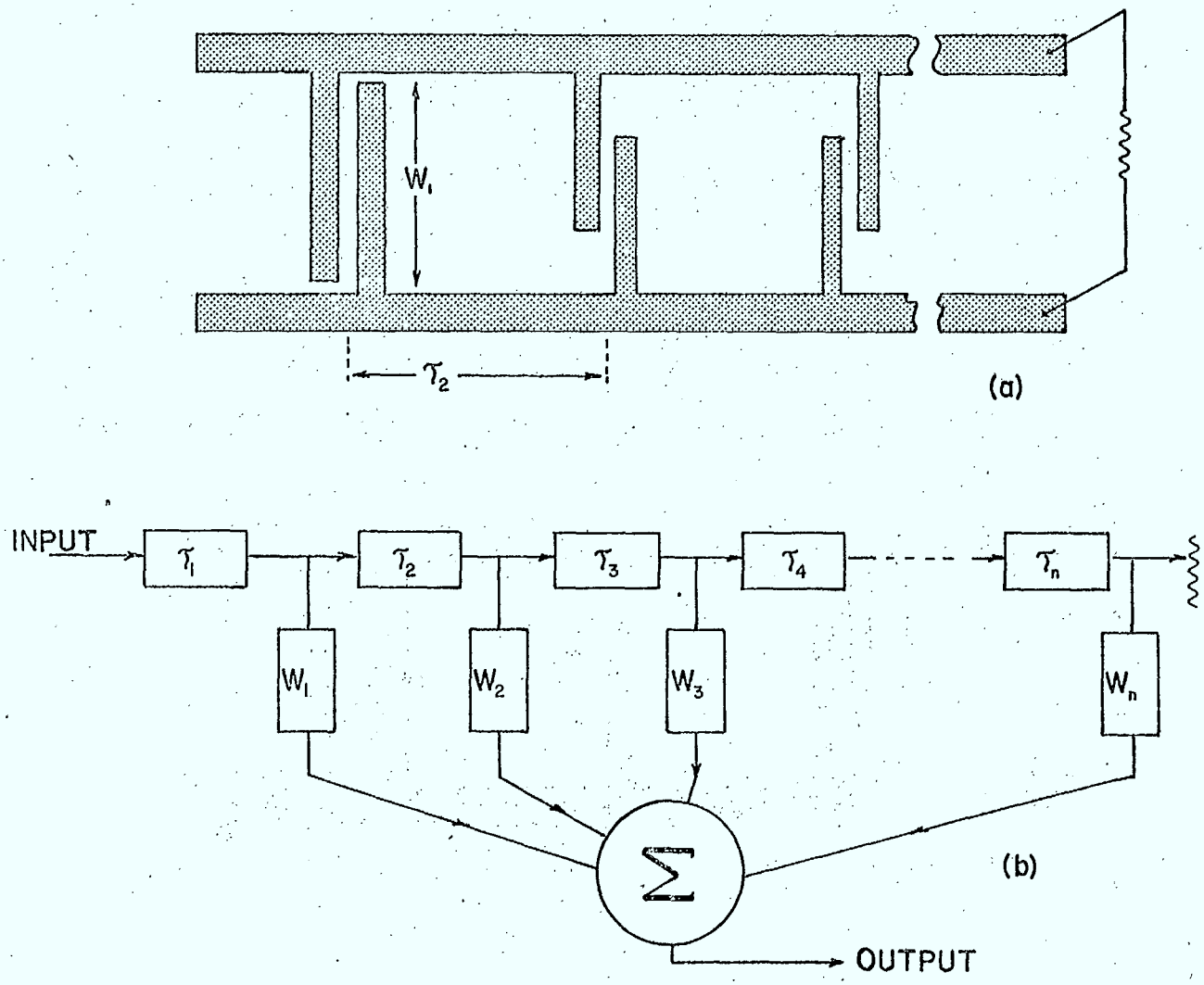


Fig. 4.1 (a) Finger pairs of an interdigital transducer.

(b) Transversal filter equivalent of the transducer in (a).

This identification with transversal filters is important because it allows so much of the previous work on analog transversal filters and current work on digital filters to be applied to the synthesis of surface-wave filters. Moreover we can note immediately that the latter filters will be of the non-minimum phase class of band-pass filters and thus the amplitude and phase responses can be specified more or less independently.

Most of the synthesis techniques for surface-wave filters are based on the time impulse response of the transducer pair, thus we should examine some of the features of the time response of surface-wave transducers. For many applications it is convenient in the synthesis and sometimes in the realization to have most of the frequency-dependent behaviour built into only one of the transducers and to make the second a uniform transducer with finger overlap and bandwidth larger than the maximum finger overlap and significant bandwidth respectively of the first transducer. This is the case which will be considered initially here and the non-uniform transducer will be taken for convenience as the input transducer. Thus we have the system block diagram illustrated in Fig. 4.2 neglecting for the moment the other acoustic port of the transducers and second-order effects. The overall transfer function will be (32)

$$H(\omega) = \frac{V_2}{V_1} = H_1(\omega)\, e^{-j\omega l/v} H_2(\omega) \qquad (4.2)$$

where here $H_2(\omega)$ will be taken as constant over the band of interest and the length l between reference points in the two transducers introduces a dispersionless delay or a linear phase shift with frequency. Thus the overall frequency response is that of the

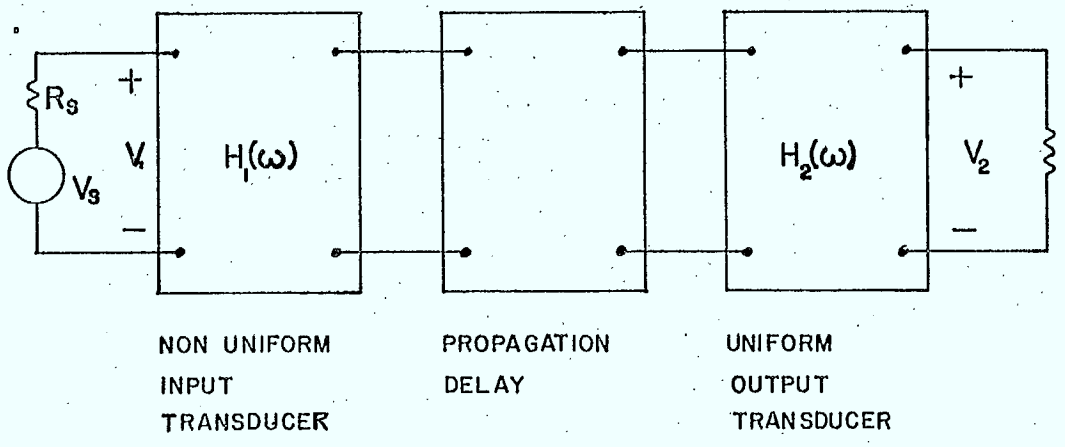


Fig. 4.2 Block diagram of surface-wave filter.

input transducer and is given by Eq. 4.1 or Fig. 4.1 using the comb as a driving rather than a receiving transducer. However for band-pass filter applications the successive fingers are all spaced by a distance of the same order and thus the length giving τ_i of Fig. 4.1 is the section width L_i of Fig. 3.4.

The steady-state frequency response of our system and its response to an impulse in time form a Fourier transform pair

$$H(\omega) = \int_{-\infty}^{\infty} h(t)\, e^{-j\omega t}\, dt \tag{4.3}$$

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H(\omega)\, e^{j\omega t}\, d\omega \tag{4.4}$$

where $H_1(\omega)$ is taken to be $H(\omega)$.

Now when a voltage impulse is applied to the transducer, electric fields of the form shown in Fig. 3.2 are set up in the vicinity of each finger at this instant of time, and as a result, a surface disturbance is launched toward the receiving transducer from each finger region and thus in the space between the transducers there will exist after the voltage impulse a wave packet which is of a length in space equal to the length of the non-uniform input transducer and which is composed of one pulse per finger pair as illustrated in Fig. 4.3a (32). The transverse width of each pulse in space is equal to the overlap length of the corresponding fingers and if the ratio a / L is the same for each finger pair the amplitude of each pulse in the wave packet will be the same except for a slowly varying correction factor to be discussed below. When this wave packet passes under the uniform output transducer it produces a train of

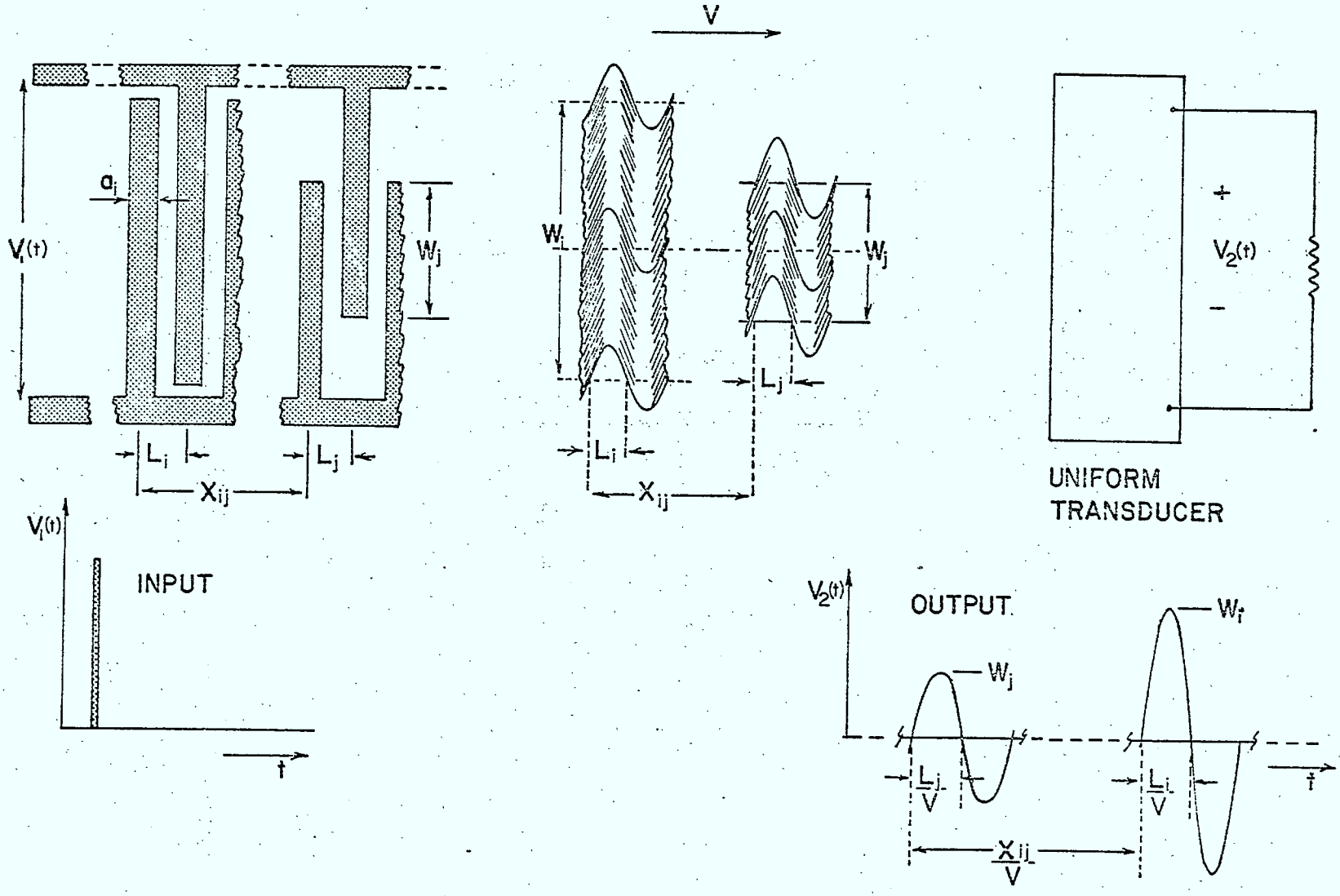


Fig. 4.3 Excitation of a surface wave by an electrical impulse with transducer of varying overlap. $v_2(t)$ is the overall impulse response of the device.

pulses each of period, spacing and amplitude-squared corresponding to the finger period, position and overlap respectively of the finger pair which produced it, as indicated by the output waveform in Fig. 4.3.

In Fig. 4.3 the individual surface-wave pulses have been sketched as individual half periods of a sine wave centred initially between two fingers. It is seen from Fig. 3.4 that the driving field distribution is very non-uniform and the resulting surface disturbance can be also expected to be of complex form (19); however, for most applications it is really only the fundamental component of this pulse and its immediate neighbours which is significant and thus using a sinusoidal pulse is a reasonable and useful approximation. This approximation and several others which have been used (such as rectangular or triangular pulses centred on or between the fingers, the sinusoid centred on a finger, or a pair of delta functions of the same sign at the two edges of each finger) all give results which predict well the fundamental behaviour but are not as valid for the harmonic responses, that is, the higher-order pass bands (11, 18, 63, 96).

The basis of a synthesis technique is suggested by the Fourier transform of Eq. 4.4 and the waveforms of Fig. 4.3. From a given frequency response the corresponding impulse response is calculated, usually by Fast Fourier Transform techniques, the impulse response is approximated, within tolerances to be discussed later, by a series of pulses of the representational shape chosen. Then a pair of fingers can be introduced into the non-uniform transducer at a relative position in space corresponding to the relative position in time of the representational pulse and of overlap length

and interelectrode spacing to produce the amplitude of that pulse. This is the fundamental concept in the synthesis of surface-wave filters but many subsidiary factors lead to different methods of applying it.

4.2 Truncation

The most important second-order consideration common to all the detailed synthesis procedures arises from the truncation of the impulse response. The impulse response of a specified bandpass filter extends over all time, but the transducer which is to provide this impulse response must be finite in length, thus the actual impulse response to be synthesized must be some finite-length approximation to the specified response. This truncation then raises the whole broad problem met in the design of analog filters, digital filters and antennas, how to produce the "best" fit to the specified response with a limited number of elements.

For band-pass filters, fitting the infinite impulse response as accurately as possible over a finite length does not necessarily produce the best approximation to the specified frequency response. For example consider an idealized rectangular band-pass characteristic with the amplitude response illustrated in Fig. 4.4 and a linear phase characteristic. The impulse response has a sin X/X form of envelope and interior oscillations approximately sinusoidal at the centre frequency. The time of occurrence of the main lobe depends on the time delay, that is the slope of the linear phase curve. Now if this impulse response is abruptly truncated in time and the Fourier transform taken, the resulting frequency response will be degraded from the original of Fig. 4.4a by the introduction of ripples in the pass band, of a finite

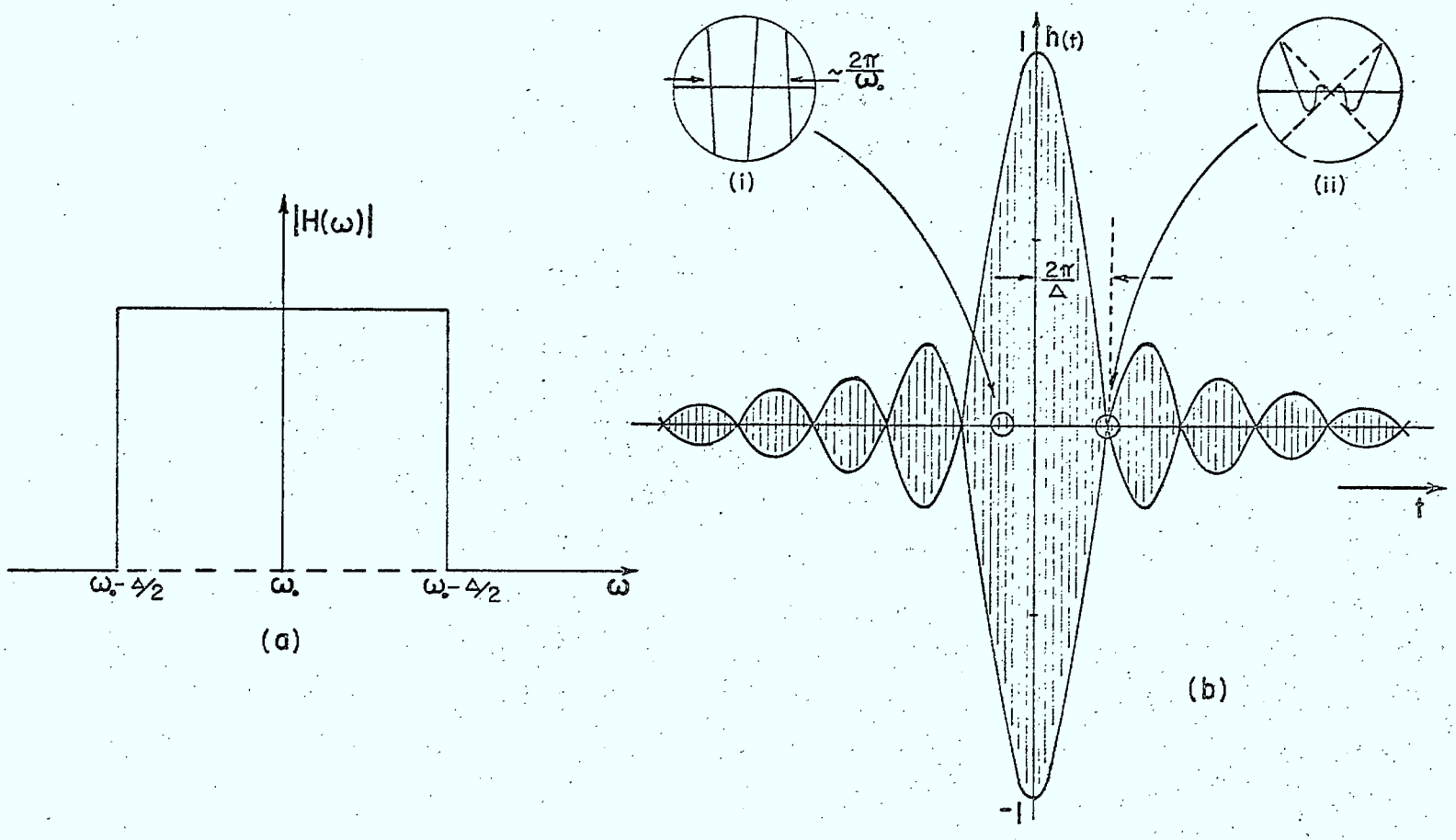


Fig. 4.4 Idealized rectangular band-pass characteristic and corresponding impulse response h (t) .

cut-off slope and of side-lobes to the main response. For example Figures 4.5a and 4.5b illustrate the band-pass characteristic which results from a symmetrical truncation of the impulse response of Fig. 4.4a at the fourth and the sixteenth zero respectively of the envelope (95). Because of the Gibb's phenomenon, only the periodicity, not the amplitude of the pass-band and stop-band ripples, is influenced by the truncation length.

Such differences between the specified frequency response and that corresponding to a truncated version of the desired impulse response leads to a requirement for iterated design procedures for the synthesis of general filter characteristics (64). That is, an initial choice of truncation function, probably more gradual in amplitude than the abrupt truncation function considered above, is applied to the desired impulse response and the resulting frequency response is recalculated and compared with the original, indicated modifications are included in the truncation function and the process is repeated. At some stage of this iteration process the transducer geometry appropriate to the truncated impulse response is included in the loop and it is the predicted transducer response calculated from as complete a model as necessary which is then used at each iteration for comparison with the specified response. Such an iteration with so many possible variables is obviously very cumbersome and the possible generality is quickly reduced by the device designer to suit the form of transfer characteristic required, the tolerances specified, available fabrication facilities and his own choice of design trade-offs. For example, in some applications it may be desirable to keep the finger spacing and width constant while varying the

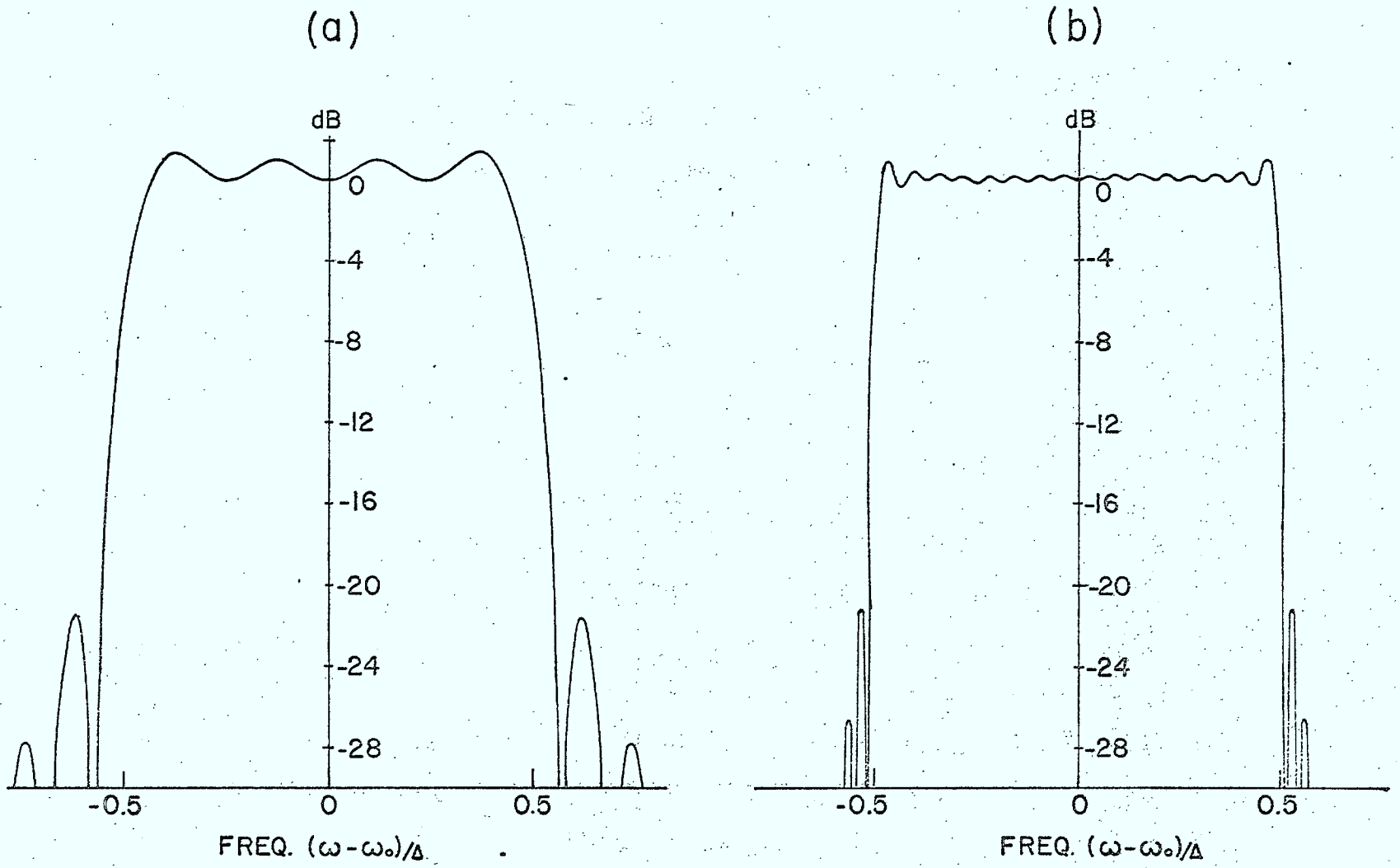


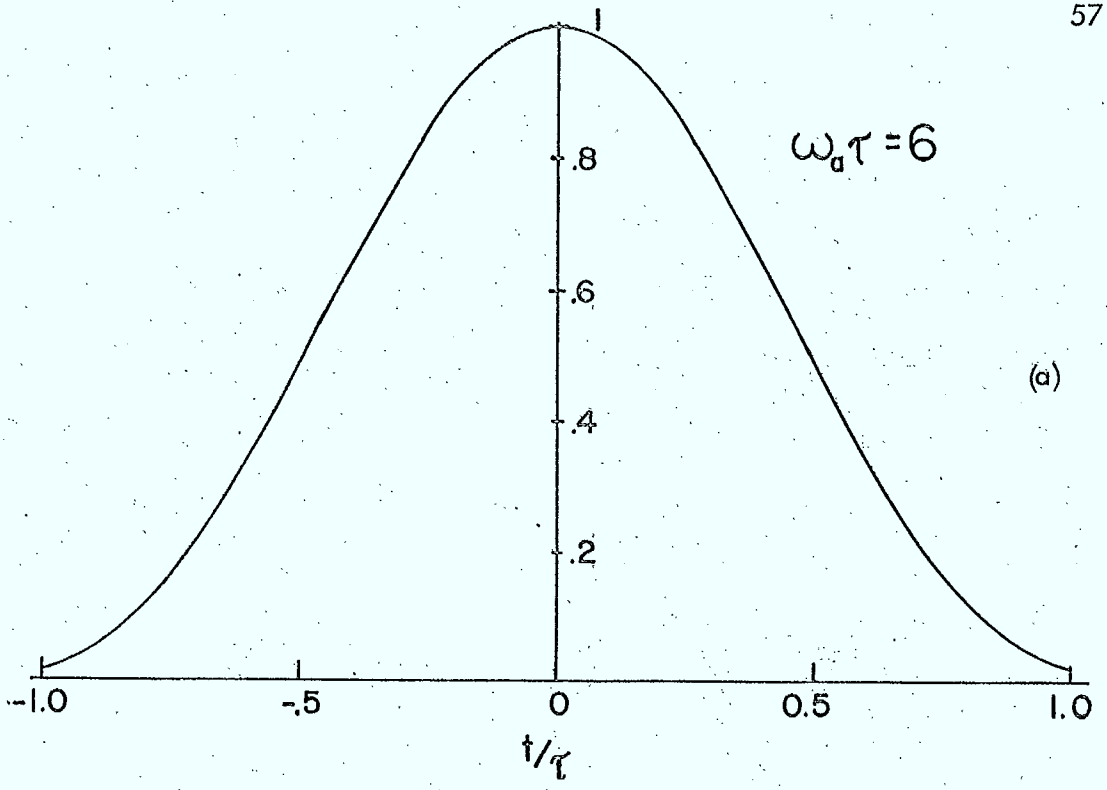
Fig. 4.5 Band-pass characteristics corresponding to the symmetrical truncation of the impulse response of Fig. 4.4b after (a) fourth and (b) sixteenth zero of envelope (95).

overlap, in other applications the overlap is held constant and the finger spacing is varied, while in yet others the overlap and spacing can be kept fixed and the inter-electrode spacing varied and in yet others the number of fixed geometry finger pairs per unit length is the defining quantity.

In any of these synthesis procedures, great use can be made of the related techniques from digital filter and antenna design. For example, instead of truncating abruptly the impulse response of Fig. 4.4b and producing the ripples in the band-pass characteristic of Fig. 4.5, the impulse response could be multiplied by one of the weighting functions used in antenna aperture weighting, such as the common Hamming and Taylor functions. Figure 4.6b shows the band-pass characteristic which results when the same length of impulse is used as in Fig. 4.5a but, with the amplitude multiplied by the Kaiser weighting function of Fig. 4.6a. This particular weighting function is given by (95)

$$w(t) = \frac{I_o\left[\omega_a \tau \left(1 - t^2/\tau^2\right)^{1/2}\right]}{I_o(\omega_a \tau)} \qquad |t| < \tau$$
$$w(t) = 0 \qquad |t| > \tau \tag{4.5}$$

where I_o is the modified zero-order Bessel function of the first kind. Table 4.1 indicates the decrease in the pass-band ripple and sidelobe level for more "tapering" of the weighting function as measured by the parameter ω_a in Eq. 4.5. It should be noted however that the tapered truncation decreases the slope of the skirts on the edges of the pass-band of the frequency characteristic; the transition width is here defined as the frequency interval in which the skirt falls from a value equal to that of the



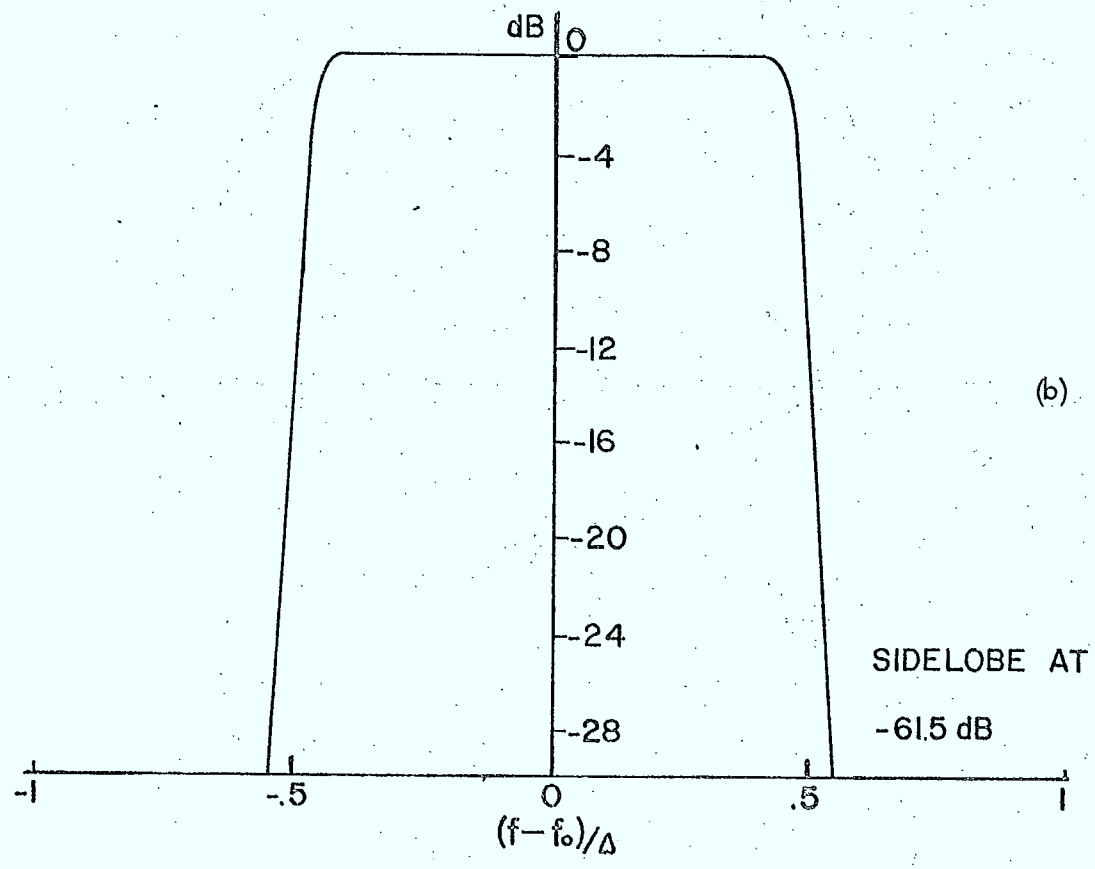


Fig. 4.6 (a) Weighting function of Eq. 4.5.

(b) Band-pass characteristic when impulse response of Fig. 4.4 is truncated by the weighting function of (a) (95) .

lowest minimum of the in-band ripple to a value equal to that of the maximum of the highest sidelobe.

Columns 2 and 4 of Table 4.1 can be used to obtain an estimate of the impulse length and hence the transducer length necessary for other forms of weighting for specified transition widths and sidelobe level. For example, a filter having 54 db sidelobes and a transition width of 2 Mhz can be expected to require a time duration of approximately 1.6 μ sec and thus a transducer length of about 5 mm on YZ Li Nb 0_3 .

It is possible to carry over to the surface-wave filter synthesis, many other results from the theory of non-recursive digital filters. As a further example, synthesizing the transversal filter of Fig. 4.1 to have an equiripple symmetrical pass-band characteristic while using equal time delays between taps gives weighting factors which can be taken as the overlap lengths of equispaced fingers in the surface-wave realization of the same characteristic. Table 4.2 illustrates some combinations of pass-band ripple and level of side lobes (all equal) for equiripple designs and the number of fingers N necessary to obtain a transition width which is 4% of the centre frequency.

To summarize this section we might note that for symmetrical band-pass filter characteristics of what might be called standard form, that is with specified centre frequency, bandwidth, pass-band ripple, out-of-band sidelobes, and skirt slope the synthesis approach is either to apply weighting functions of known properties to the rectangular band-pass impulse response or to use the iterative techniques of

TABLE 4.1

TRANSITION WIDTH (TW) AND RIPPLE FOR A BESSEL WEIGHTING FUNCTION OF TIME DURATION 2τ (95)

Parameter $\omega_a \tau$	FILTER RESPONSE		
	TW x τ	Peak-to-Peak Ripple at Band Edge (dB)	Peak Out of Band Sidelobe (dB)
0	0.50	1.4	-21
1	0.56	1.0	-24
2	0.80	0.52	-30
3	1.03	0.21	-37
4	1.32	0.076	-45
5	1.64	0.026	-54
6	2.00	0.0094	-62
7	2.23	0.0040	-71

TABLE 4.2

RIPPLE AND TRANSITION WIDTH FOR AN EQUIRIPPLE FILTER WITH (2N + 1) ELEMENTS (95)

Peak-to-Peak Passband Ripple (dB)	Stopband Ripple (dB)	Parameter $N \times \frac{\text{Transition Width}}{\text{Center Frequency}}$	Approximate Number of Elements (2N + 1) for $(\frac{TW}{F_o}) = 0.04$
0.55	- 40	1.9	96
0.55	- 50	2.0	103
0.55	- 60	2.2	110
0.17	- 40	2.0	101
0.17	- 50	2.2	115
0.17	- 60	2.5	129
0.055	- 40	2.1	106
0.055	- 50	2.5	127
0.055	- 60	2.9	147

digital filter design (64) for say equiripple pass-bands to produce the truncated impulse response. For less standard filter characteristics, e.g. a T.V. I F filter, the impulse response to be synthesized must be truncated and weighted in a much more ad hoc fashion with iterative attempts at the weighting impulse length and sample locations and with no guarantee of convergence nor easily specified optimum result (16). Since we are dealing here with surface-wave device principles let us leave this important problem of the selection of an optimum impulse response for a specified filter characteristic which is a common problem with other forms of filter synthesis and return to some of the specific problems of the realization of the desired impulse response by surface-wave transducers.

4.3 Finger Spacing and Weighting

Continuing to assume that one transducer is broad-band and uniform and that the desired impulse response is built into the other, the fingers of the latter will be placed at positions in space corresponding to times at which the impulse response is a local maximum or the interelectrode gap is located at positions corresponding to the zeros in Fig. 4.4b. If the response crosses the axis as is usually the case, inset (i) of Fig. 4.4b, then adjacent fingers are attached to opposite bus bars, whereas at phase reversal points such as shown in the inset (ii) of Fig. 4.4b adjacent fingers are attached to the same bus bar. There is negligible modification of the overall response if the fingers are centred on the zeroes and this is sometimes easier to program.

The "weight" to be produced by a given finger is the local amplitude of the desired impulse response, that is the local value of the envelope function (e.g. Fig.4.4).

Providing there are no abrupt discontinuities in the local pattern the contribution of a single finger to the overall impulse response w_i is proportional to three factors that is

$$w_i = \text{const } W_i S_i L_i^{-3/2} \tag{4.6}$$

where W_i is the overlap length with adjacent fingers, S_i is a factor dependent on the metallization ratio and is given approximately by the square root of the term in square brackets of Eq. 3.2 or more directly by (G4, p.134)

$$S_i \cong 0.37 + 1.69\, a_i/L_i - 0.91\, (a_i/L_i)^2 \tag{4.7}$$

where a_i is the finger width and L_i the interelectrode distance (Fig. 4.3). S_i arises because of the change in the field distribution with metallization ratio and the $L_i^{-3/2}$ dependence arises from frequency scaling rules (26). The range of finger "weights" available through S_i alone (19) is limited by tolerances on gap and finger widths to a ratio of about 3 : 1 and thus for most applications it is most convenient for fabrication tolerance to keep the ratio $a_i/L_i = 1/2$ and to realize the impulse response by variations in finger position and overlap length only. The slowly varying factor $L_i^{-3/2}$ need be included only if there is appreciable change in finger spacing along the length of the transducer. Similarly it is sometimes more practical not to make the other uniform transducer so much broader than the weighted one that it does not influence the characteristic of the latter but rather to include compensation for its roll-off in frequency in the initial frequency specification before the design impulse response is fixed.

It can be easily shown from Eq. 4.1 that if the transducer tap weighting function w_i of Eq. 4.6 is symmetric or antisymmetric about the central line of the transducer, the frequency response has linear phase with frequency, that is, it is dispersionless (51). Moreover the phase factor at the acoustic output port of the transducer corresponds to the time required to travel from the centre line of the transducer to this port and thus for overall phase calculations it can be assumed that the source of the acoustic wave emanating from a symmetrical or anti-symmetrical transducer is the central line. This ease of obtaining linear phase is often one of the major advantages of surface-wave over conventional filter designs. On the other hand shaped phase curves are often required and some non-symmetrical transducers with dispersion are considered in Chapter V.

The input impedance of a weighted transducer can be calculated by interconnecting individual equivalent circuits of the form of Fig. 3.4 or Fig. 3.8, one for each finger with parameters suitably chosen for that finger section. It is not possible to obtain a simple expression of the form of Eq. 3.5 or Eq. 3.6 and accurate calculations for weighted transducers require a computer program (63, 81, 97). However the input impedance will consist of the static capacitance, which can be determined by adding the finger section capacities given by Eq. 3.2, in parallel with a radiation susceptance which can often be neglected, and a radiation conductance whose frequency dependence is proportional to the magnitude squared of the transducer frequency response, $|H(\omega)|^2$. Estimates of input admittance can be made from the curves of Fig. 4.7 (26) which show the approximate input conductance (G) and

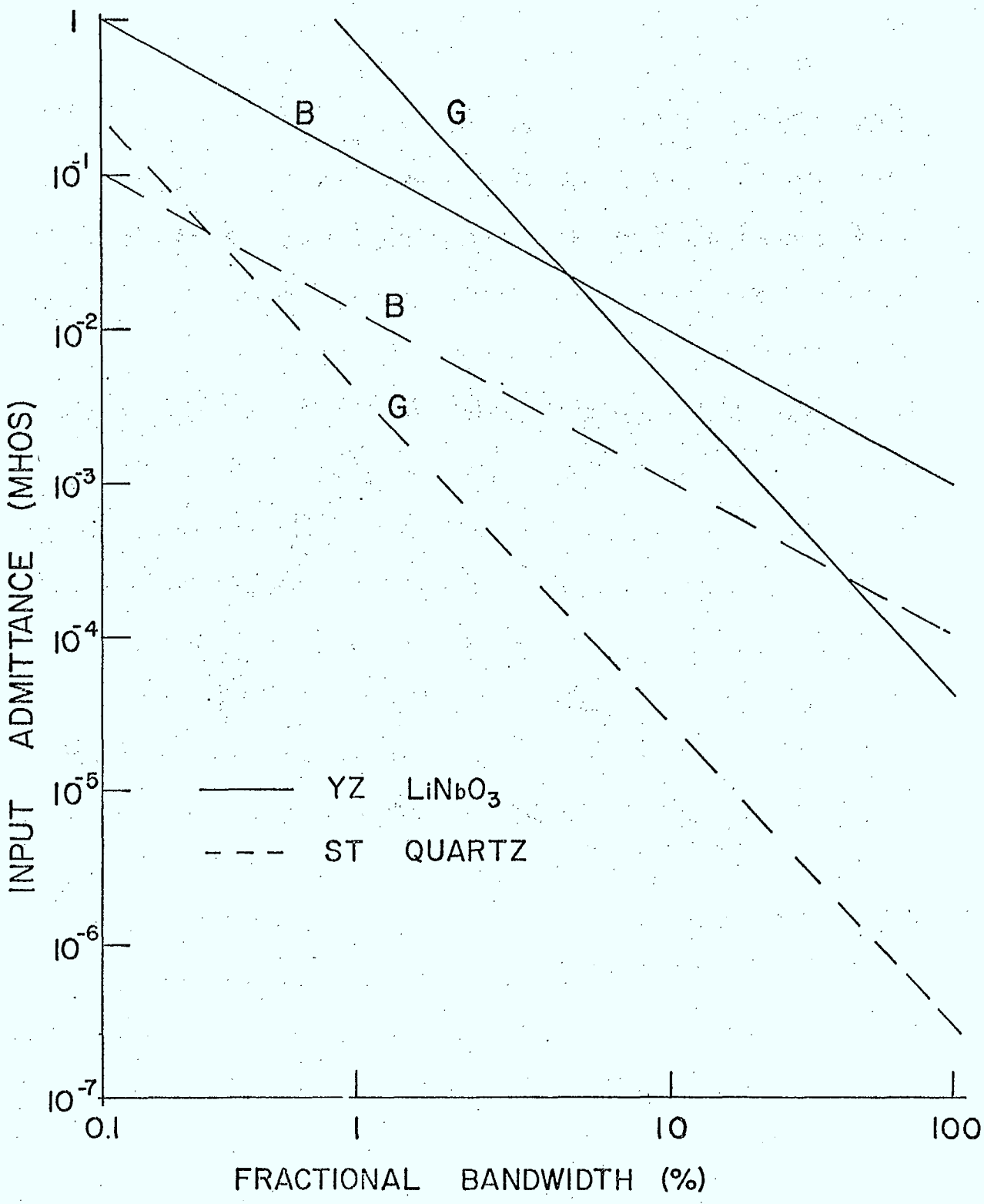


Fig. 4.7 Input admittance for approximately rectangular band-pass characteristics for an assumed average beam width of 100 λ .

susceptance (B) for essentially rectangular band-pass filters on two different substrates as a function of the fractional bandwidth for an assumed average beam width of 100λ. These values can be scaled proportionately for beam width and if the transducer is dispersive the values shown are to be multiplied by the time-bandwidth product (see Chapter V).

4.4 Apodization Corrections

There are two effects which arise with the finger-length weighting or apodizing which do not occur with constant-length fingers. The first of these effects can be visualized with the help of Fig. 4.3. The part of the wave from the long set of fingers which is excited near the longitudinal centre line of the transducer passes under more fingers than the part of the same wave excited near the finger extremities because some of the other fingers in the transmission path are shorter. Since the phase velocity in the metallized region of a piezoelectric is lower than in a non-metallized region, the average phase velocity of the part of the wave excited at the extremities of the long fingers in Fig. 4.3 is lower than that of the central part of the wave, serious phase distortion will occur across the wavefront arriving at the receiving transducer, especially if the non-uniform transducer is long and has a wide range of finger lengths. This phase distortion could be included in the weighting but only with great difficulty because the weight of any one finger pair depends on the average metallization of the others. However in fabrication this effect can be removed very simply by using dummy fingers (21), (G6, p.419) as illustrated in Fig. 4.8. Since each of the dummy fingers introduced is connected to the same bus

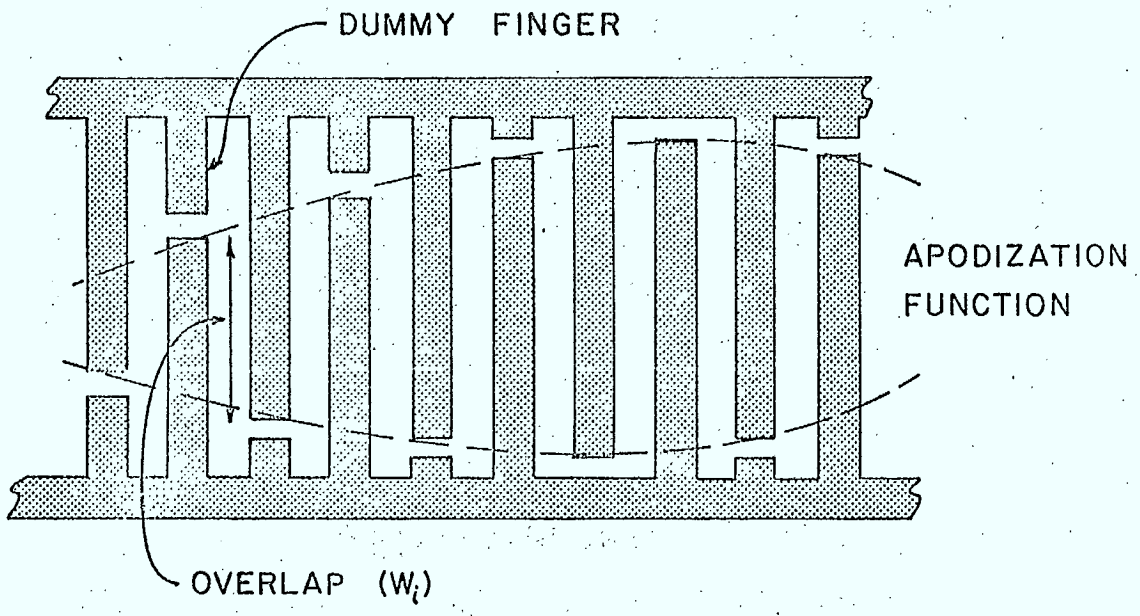


Fig. 4.8 Use of dummy electrodes to compensate for phase errors in apodized transducers.

bar as its immediate neighbours no new electric fields are set up by it and thus the weight of each finger pair remains the same, but now any part of a wave propagating under the transducer sees the same average metallization.

The second effect associated with variation of finger overlap length can arise at the output transducer. At first sight from Eq. 4.2 it might appear that a given transfer function could be realized by dividing it into a product of two and realizing the two individual functions by the overlap weighting of two transducers in the manner discussed to this point. But the latter approach has assumed that all of the wave emitted by a finger pair of the transmitting transducer passes under all of the length of all of the fingers of the receiving transducer. However if both transducers are apodized the part of the wave emitted by the extremities of long fingers will not see the short fingers of the receiving transducer. Thus the overall transfer function of such a pair will not in general be the product of the individual transfer functions as synthesized from the impulse response illustrated in Fig. 4.3 (63). Pairs of transducers in which both transducers are apodized can be analyzed by imagining each transducer and the transmission path between to be divided into narrow strips parallel to the central axis, within each strip all the finger pairs of each transducer have uniform overlap equal to the strip width and thus the contribution to the impulse response produced by the strip can be calculated and addition of the responses of all of the strips gives the total response of the transducer pairs (96). This additional complication in the synthesis is inconvenient and if the necessary filter characteristic, e.g. a steep skirt requirement, cannot be met by one apodized

transducer, several techniques can be used to circumvent the strip analysis and retain the overall frequency response as the product of the individual ones. For example the width weighted acoustic beam can be converted to an amplitude-weighted beam of uniform width by means of a multi-strip coupler (Chapter IX) between the transducers (G6, p.172). Or monotonic variation of the finger spacing along the transducer length can be used with two identical transducers so that the wave emitted by fingers of a certain spatial frequency is received in large part by the fingers of the same spacing and length in the receiving transducer. As will be seen below the magnitude of the overall frequency response is the square of the individual frequency responses and the overall phase may be linear or dispersive, linear if the second transducer is a translation of the first. (Chapter V).

Another method of tap weighting is illustrated in Fig. 4.9 (G4, p.141) where it will be noted that the generated wave is essentially uniform over the full aperture width. This method is difficult to use alone for complicated impulse responses but can be used in conjunction with overlap weighting of the individual generating region to produce a more-or-less uniform beam over the aperture and to avoid the very short fingers which lead to extreme diffraction effects in normal apodizing for some impulse responses.

There are other methods for varying the frequency response of a pair of transducers while maintaining constant beam width. Withdrawal weighting (G6, p.423) has been used in which the number of constant amplitude samples per unit time of the impulse response is varied in a manner analogous to pulse-number modulation. Yet

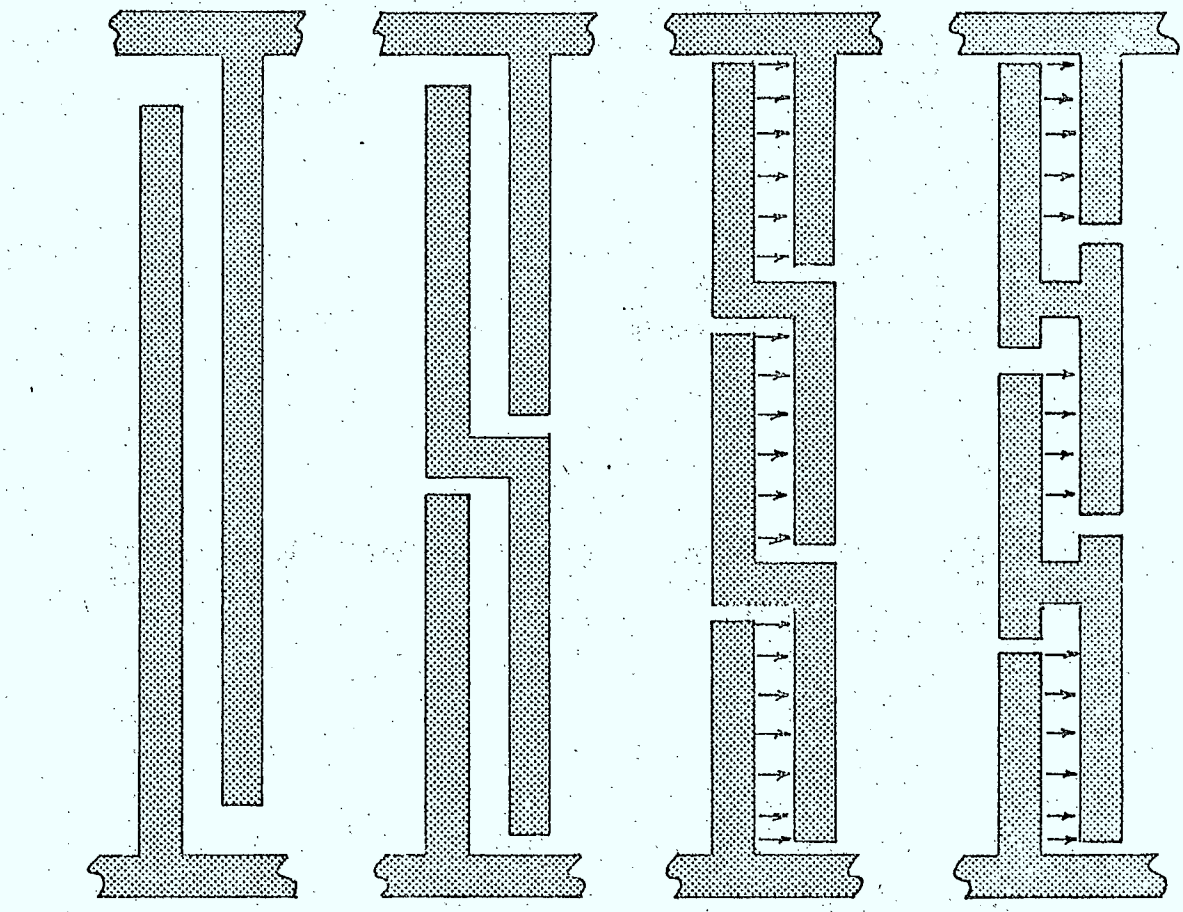


Fig. 4.9 Finger overlap weighting by series connection of electrodes (capacitive coupling).

another constant-beam-width method which has some advantages when the location and depth of nulls in a response are important is the phase weighting technique illustrated in Fig. 4.10. Since the overall transducer is symmetric it is non-dispersive (Section 4.3) and will appear to the receiving transducer as a source located on the central line marked 0. This particular transducer can thus be regarded as two separate sources located on the central line, one with a frequency response characteristic of the central seven fingers and the other with a frequency response of the outer three-finger transducers. However, in adding the sources together, the relative phase of the two sources $\omega X/v$ must be taken into account in the vector addition. The resultant frequency response of such a symmetric transducer can be asymmetric about the centre frequency; for example in Fig. 4.10 if both transducers have the same centre frequency and at this centre frequency X corresponds to say $(n + \frac{1}{4}) 2\pi$, the resultant will be larger at frequencies below the central frequency than above. It does not appear that there is a straightforward method of using this frequency-domain analysis in a direct synthesis technique. That is, while iteration is used to improve the initial design in the time-domain synthesis techniques previously discussed the initial configuration was reasonably well defined, here the initial configuration choice for an iteration procedure is much more arbitrary. Thus phase weighting has the advantage of dispersionless response and constant finger overlap, but how wide its use becomes in surface-wave filters will depend on the refinement of the associated synthesis procedures.

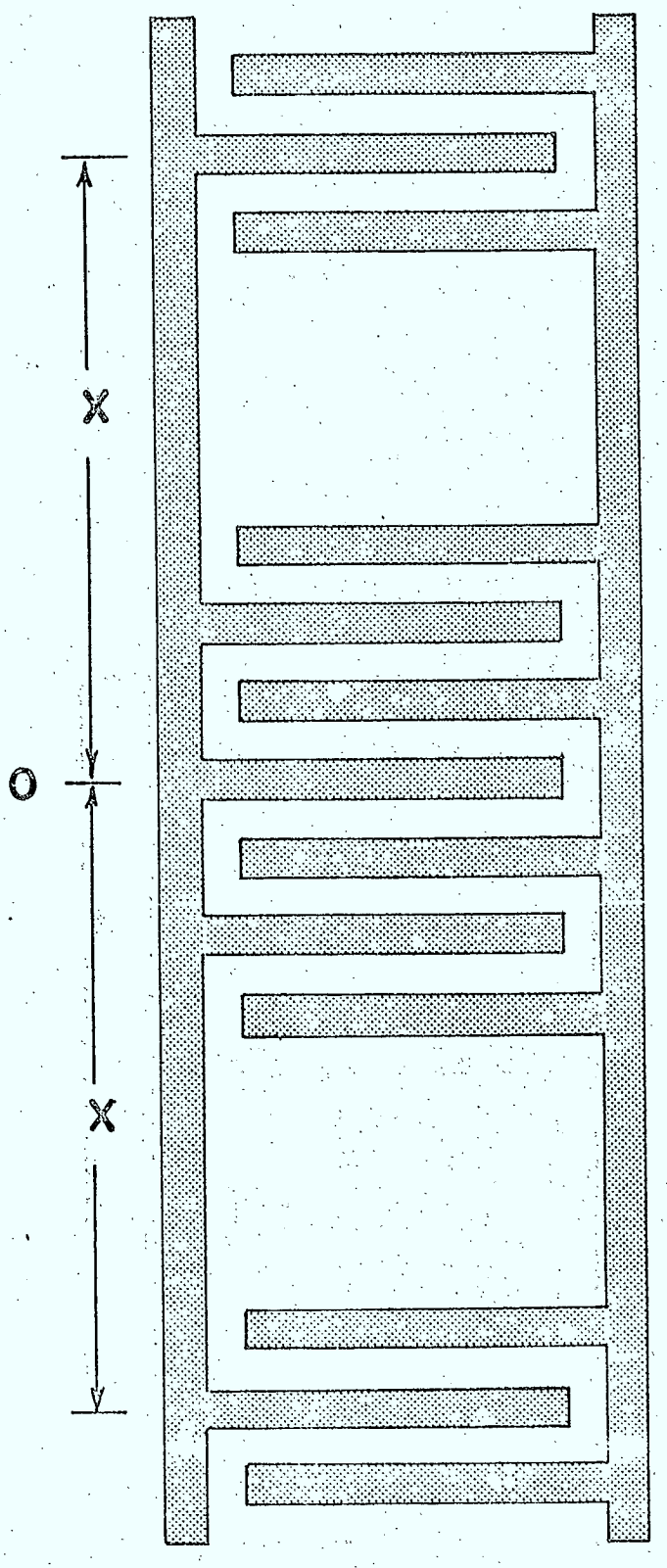


Fig. 4.10 Phase-weighting for non-dispersive band-pass filters.

Finally, it should be noted here again that in all the transducers weighted to produce specified filter characteristics triple transit echoes between the two transducers (Section 3.4) can deteriorate the frequency response from its design value, especially on high coupling substrates. Such echoes can be greatly reduced by using the split fingers of Fig. 3.9 for each of the fingers of each transducer and will only negligibly affect the filter design for the fundamental frequency band because in such a design the exact shape of the field distribution in the finger region only affects the scale factor of the overall response.

V. LINEAR FM (CHIRP) FILTERS

There is one particular type of band-pass filter characteristic which is easily realized with surface-waves and which is of importance in radar applications. The characteristic in question has a time response which is a constant amplitude RF pulse whose instantaneous frequency during the pulse varies linearly with time, or whose instantaneous phase varies quadratically with time.

In many radar systems, the peak radiated power and thus the maximum range is limited by the transmitter. Long pulses are required to give adequate echo energy, but long pulses have inherently poor range resolution unless some form of coded detection is used. One standard technique is to "chirp" the outgoing pulse, that is to vary the frequency radiated in a linear fashion over the pulse duration. If the returned echo is converted into a surface-wave by a broad-band transducer and passed under a second transducer whose length corresponds to the pulse length and whose finger spacing varies linearly with distance along the transducer in a time-reversed fashion from the instantaneous frequency variation of the transmitted pulse, then the transducer forms the matched filter of the transmitted pulse and an output correlation voltage pulse appears across the transducer when the travelling pulse is precisely under it. The narrow output pulse produced by matched filtering allows high range resolution, that is, the ability to distinguish between two targets closely spaced in range is enhanced without decreasing the detecting range of the system. The surface-wave chirp filter will be discussed in more detail below but first a related signal-processing application of great potential importance for the same form of chirp filter should be noted.

Let us consider the operation of a system outlined in the block diagram of Fig. 5.1a where the linear FM compressor has a time delay vs frequency characteristic as illustrated in Fig. 5.1b and the local oscillator is scanned linearly from a frequency f_a to a lower frequency f_b. For illustrative purposes assume that two pulses of length equal to the total time delay of the compressor but of different frequencies f_1 and f_2 arrive simultaneously at the input A, then at the output of the modulator there will be two simultaneous pulses, one scanned from $f_1 + f_a$ to $f_1 + f_b$ and the other from $f_2 + f_a$ to $f_2 + f_b$. The compressed pulses from each of these simultaneous input pulses will appear at the output at a different time (G6, p.494). Thus for each frequency component in the input signal sample there is an output pulse of proportional amplitude and displaced in time according to that input frequency. For example if the input signal is a sample of some video signal, the envelope of the compressed pulse is the amplitude of the frequency spectrum. The output can be returned to base band by a second linear frequency modulation from f_b to f_a. Use of two parallel systems with quadrature phase on the scanned local oscillator are required to produce both the real and imaginary parts necessary for say a Fourier transform of a sample of an arbitrary input signal. Extension of the concept Fig. 5.1 to systems applications is beyond the scope of this report; but it may be remarked that because of the convenience illustrated below with which surface-wave FM pulse compressors can be realized, such spectral analysis applications could well become the most important commercial domain for surface-wave devices (22), (G4, p.167 and 278).

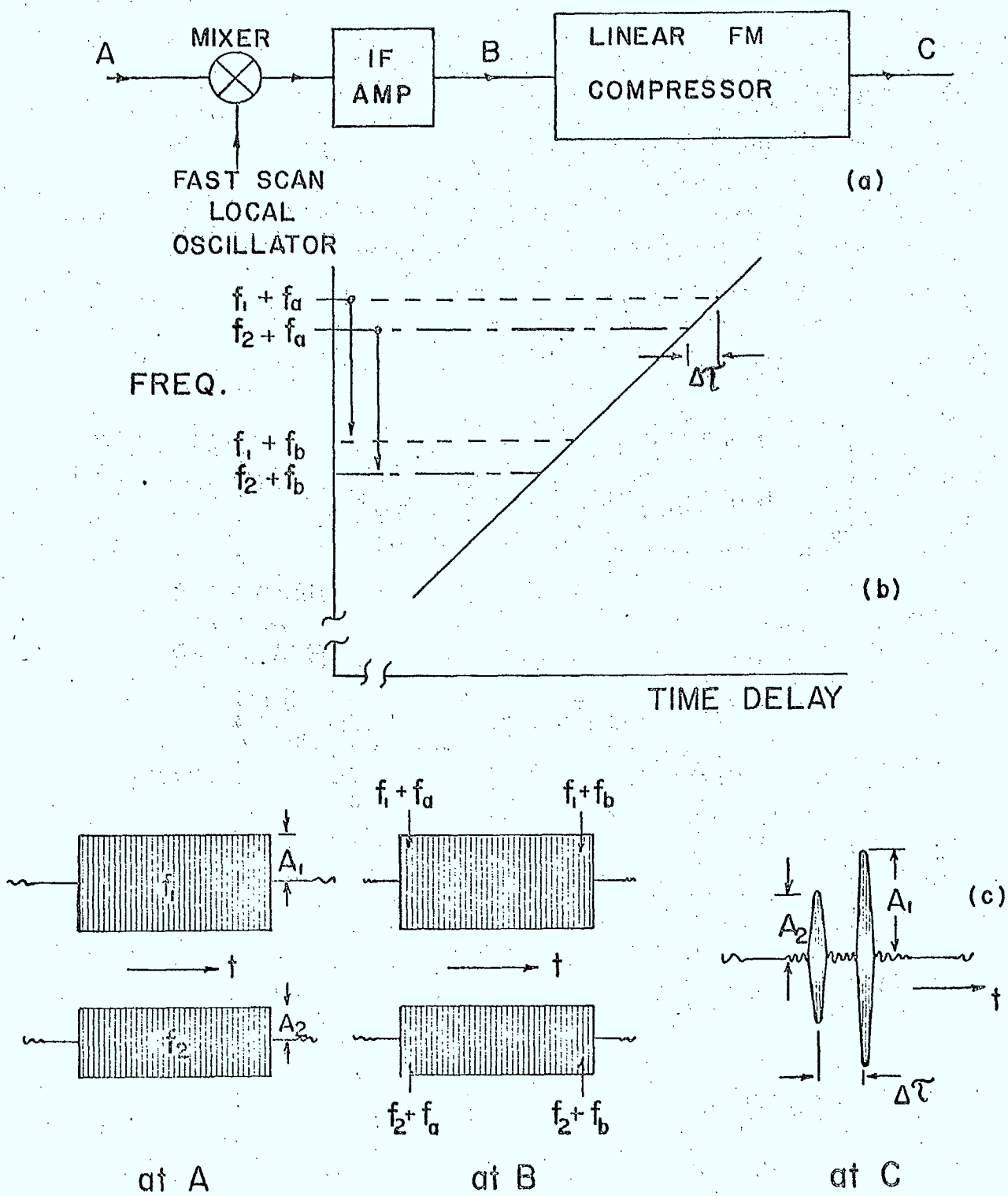


Fig. 5.1 Time separation of overlapping pulses of different frequencies. Oscillator scans from f_a to f_b during the pulse interval. (b) is dispersion characteristic of chirp filter.

5.1 Chirp Filter Design

We will now consider some design aspects of the surface-wave realization of the linear-chirp filters. The impulse response of a transducer can be written as

$$h(t) = A(t)\, e^{j\varphi(t)} \tag{5.1}$$

where $A(t)$ and $\varphi(t)$ are real functions of time and for the band-pass types of response of concern here $A(t)$ can be taken as a slowly varying amplitude or envelope function on the complex exponential which in turn has a slowly varying phase. In Fig. 4.4 for example $A(t)$ was of $\sin X/X$ form and $\varphi(t)$ was a linear $f_o t$, that is the instantaneous frequency

$$f = \frac{1}{2\pi} \frac{d\varphi(t)}{dt} \tag{5.2}$$

was a constant.

In a chirp filter the basic goal is to produce an impulse response in which $A(t)$ is constant over a pulse length T and the instantaneous frequency varies linearly with time, from $f_o - \Delta f/2$ at the beginning of the pulse to $f_o + \Delta f/2$ at the end (or vice versa) that is

$$A(t) = \begin{cases} 1 & -T/2, < t < T/2 \\ 0 & |t| > T \end{cases} \tag{5.3}$$

and $$f = f_o + \frac{\Delta f}{T} t \tag{5.4}$$

or $$\varphi = 2\pi \left(f_o t + \frac{1}{2} \frac{\Delta f}{T} t^2\right) \tag{5.5}$$

to within a constant. In principle the same transducer configuration can be used with a short electrical pulse input to produce the expanded linear FM pulse as is used to compress the linear FM pulse back to an output pulse of short duration. However as will be noted in an example below, system needs may make certain modifications desirable.

Two prototype configurations for linear FM chirp filters are illustrated in Fig. 5.2 (G4, p.167). If in Fig. 5.2a the left-hand transducer is uniform and has a bandwidth wide compared to the chirp bandwidth Δf, then all frequencies effectively radiate from its centre line (Section 4.3) so that if the finger spacing of the long right-hand transducer varies from $L_i = v/2\ (f_o + \Delta f/2)$ at one end to $L_i = v/2\ (f_o - \Delta f/2)$ at the other the configuration is matched to an "up" chirp. The higher frequencies are "resonant" at the left hand end of second transducer and are extracted there while lower frequencies proceed further through the transducer before becoming resonant with the local finger spacing. Thus when an up-chirped linear FM pulse is applied to the input transducer the low frequency part of the signal which is emitted first can have the same time delay to the output circuit as the high frequency components which are emitted later but travel a shorter distance along the substrate. Thus the output voltage pulse receives an integrated (correlated) contribution from all parts of the input pulse but is much shorter than the latter. Figure 5.2b shows a second configuration for the same type of pulse compression but here both transducers are dispersive one being the mirror image of the other

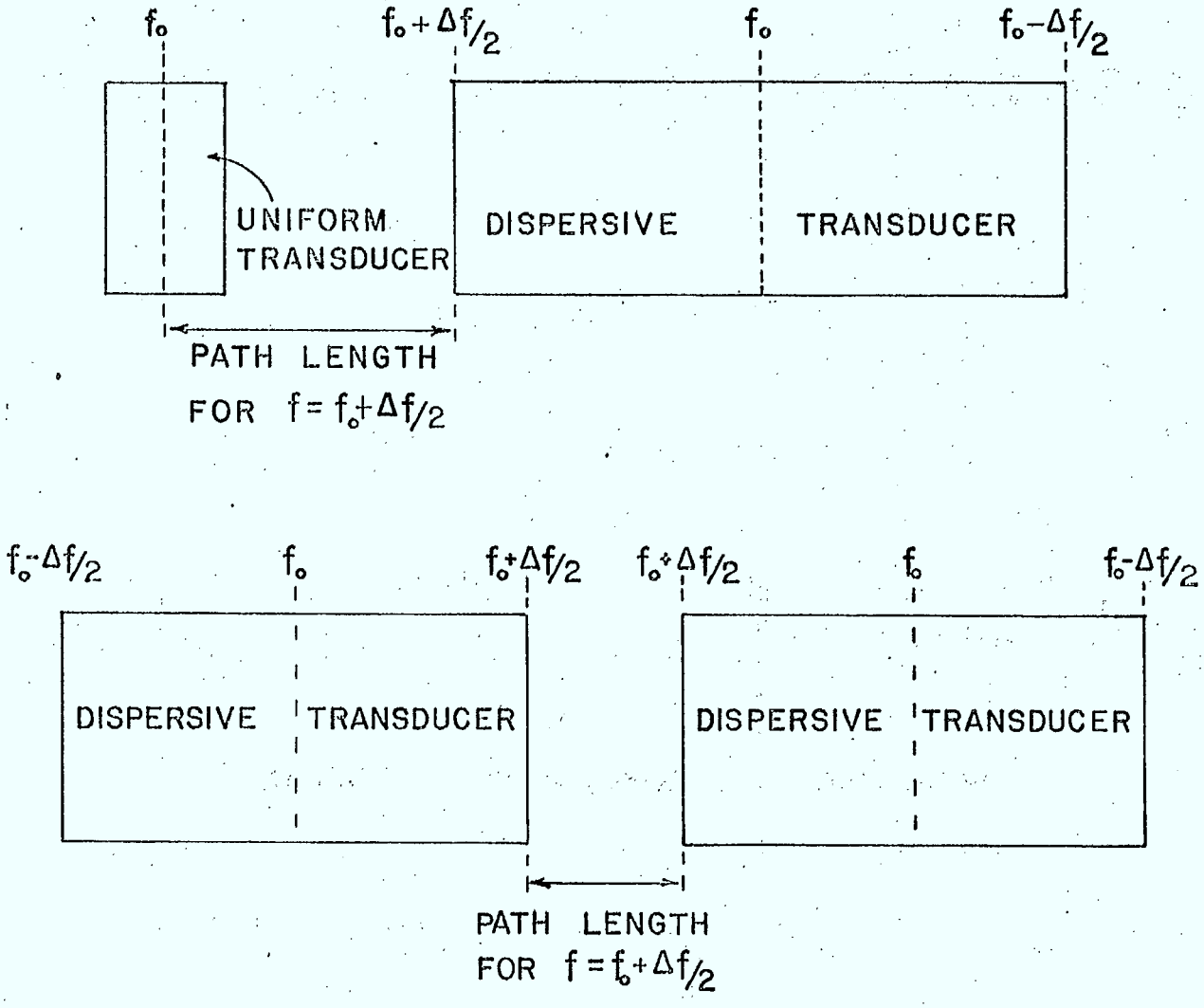


Fig. 5.2 Pulse compressors using one or two dispersive transducers.

In a design context the centre frequency f_o, the bandwidth Δf and the expanded pulse length T are usually determined by system constraints. The surface-wave filter is to approximate the phase characteristic of Eq. 5.5 as well as possible over the time interval $-T/2 \leq t \leq T/2$. It was seen in Section 4.1 that the fingers of an interdigital transducer whould be placed at the zero crossings of the desired impulse response (87, 96), so that here the fingers are to be located at $x_n = v\, t_n$ where

$$\varphi(t_n) = n\pi \qquad n = 0, \pm 1, \ldots \tag{5.6}$$

and using Equation 5.5

$$t_n = f_o \frac{T}{\Delta f} \left[-1 + (1 + n\ \Delta f / T f_o^2)^{1/2}\right]. \tag{5.7}$$

The limits on n are $T(f_o + \Delta f/4)$ on the high frequency side and $T(f_o - \Delta f/4)$ on the low giving a total number of fingers

$$2N = 2 f_o T . \tag{5.8}$$

For several considerations such as the specification of matching networks it is helpful to have an estimate of the number of fingers in the transducer which are effective at any one frequency, that is how many fingers would a similar uniform transducer need to have in order to produce at a given frequency the same CW response as the chirped transducer (88). The estimate can be made by noting that for a specified frequency in the pass-band a certain point in the transducer will have resonant spacing and

contribute in phase while some distance on either side the phase of the fingers will have shifted by more than $\pi / 2$. Using the $\pi / 2$ phase as an estimate of the effective number of fingers N_e at the frequency f gives

$$N_e = 2 f \sqrt{T / \Delta f} \quad . \tag{5.9}$$

5.2 Filter Example

It is instructive to consider a specific example of a surface-wave chirp filter (G6, p.474). Let us assume:

$$f_o = 60 \text{ MHz}, \qquad \Delta f = 20 \text{ MHz}, \qquad T = 2 \mu \text{ sec}.$$

The desired instantaneous frequency and phase,

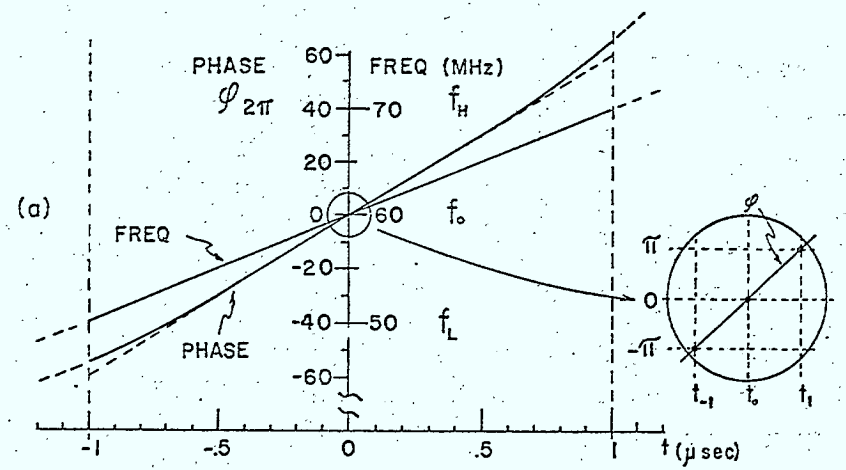
$$f = 60 + 10\, t \text{ Mhz}, \qquad \varphi = 2\pi\, (60\, t + 5\, t^2)$$

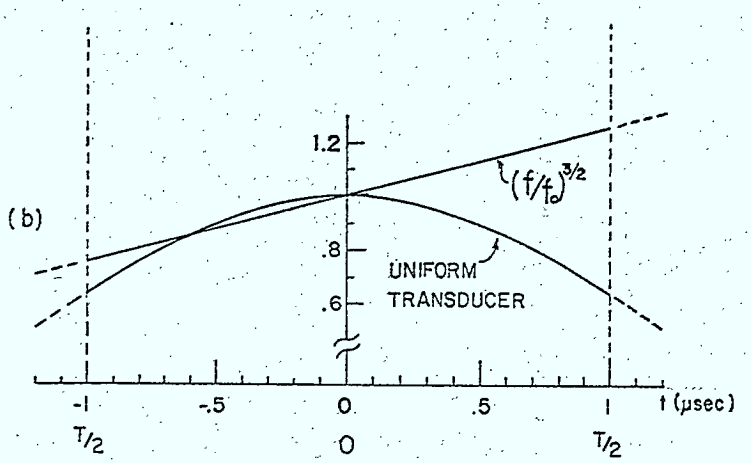
respectively with t in microseconds are shown by the solid curves in Fig. 5.3, and the time positions of the fingers

$$t_n = 6 \left[-1 + \sqrt{1 + n / 360} \right] \mu \text{ sec}$$

are illustrated by the inset diagram. The required number of fingers, Eq. 5.8, is 240 with $-110 \le n \le 130$. At the centre frequency the effective number of fingers is 38, and the finger spacing in time is $1 / (2 \times 60 \text{ MHz}) = 7.14$ n sec. at the high frequency end and 10 n sec at the low frequency end. The actual finger positions $x_n = v\, t_n$ depend on the surface-wave velocity for the chosen substrate.

If the fingers in this example all have the same ratio of metallization to finger spacing then the overlap length must be multiplied by a factor $(f / f_o)^{-3/2}$,





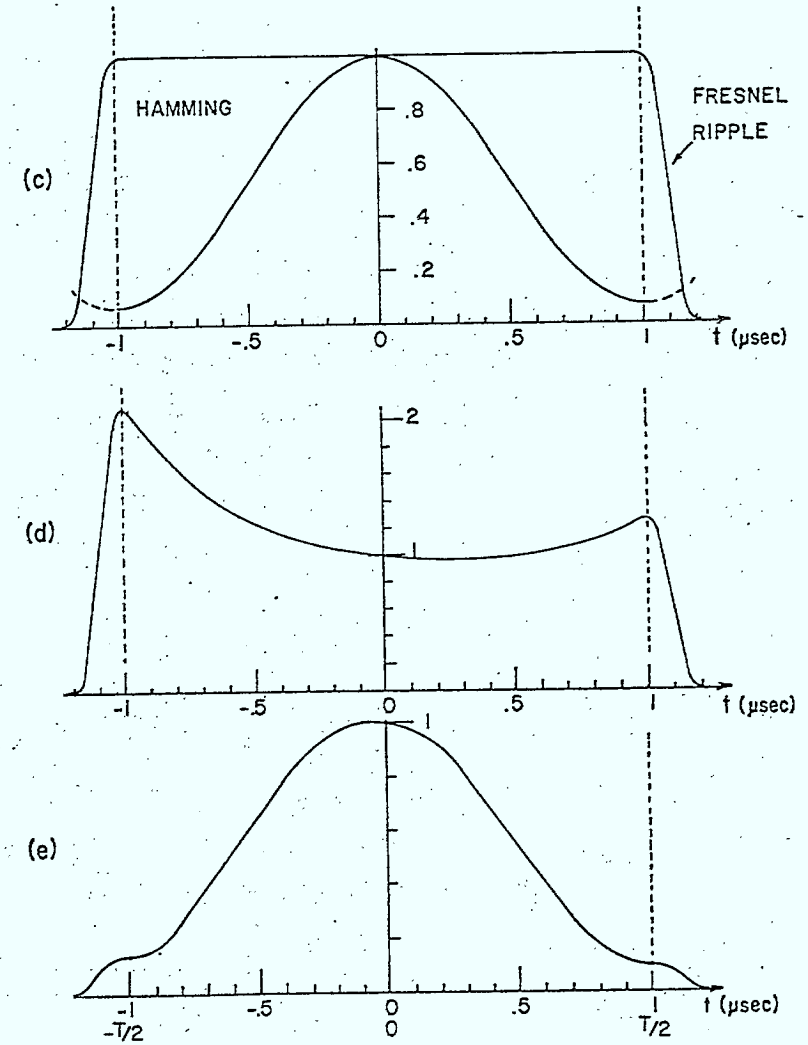


Fig. 5.3 (a) Desired frequency and phase characteristic of filter.

(b) Response of uniform transducer and excitation characteristic of fingers of non-uniform transducer, translated to a time dependence.

(c) Weighting functions for Fresnel ripple and time sidelobes.

(d) Apodizing function for expondor.

(e) Apodizing function for compressor.

Eq. 4.6, to maintain a constant response with changing spacing. Moreover, in order to have a constant overall response during the pulse it is necessary to compensate for any roll off in the frequency response of the uniform transducer over the frequency band Δf. The top of a $\sin X / X$ response for a three-finger-pair (six finger) uniform transducer is plotted in Fig. 5.3b as a function of the time in the pulse at which the different frequencies occur. Also shown in Fig. 5.3b is a curve corresponding to $(f / f_o)^{3/2}$, thus to produce uniform response over the pulse length the transducer finger lengths should be weighted or apodized according to the reciprocal of the product of these two curves. Thus the required overlap length as a function of finger length in the interval $-T/2 \leq t \leq T/2$ is as shown in Fig. 5.3d.

Abrupt termination of the transducer at $|t| = T/2$ would result in the phase and amplitude vs time curves, that is the impulse response, having ripples analogous to the frequency response ripples produced by the truncation of the non-dispersive transducers of Section 4.2. Here the resulting ripples are usually referred to as Fresnel ripples because of the similarity to the spatial Fresnel diffraction from a limited radiating aperture. These ripples in the phase characteristic can be reduced by tapering the apodizing function gradually to zero outside of the fundamental length. A cosine taper adding 20% to the transducer length is shown in Fig. 5.3d. Thus the transducer pair, one broad band and one dispersive with this apodization function, produce an overall impulse response which is a linearly chirped f m pulse of constant amplitude over the central 2 μ sec. The main departures from ideal response result from triple transit echoes, interelectrode coupling in the dispersive transducer, and

bulk-wave generation; these can be reduced by the use of split electrodes, multi-strip couplers and bulk-mode suppressors, but most easily by weak coupling to the substrate if the additional insertion loss can be tolerated. Departures from quadratic phase of perhaps $\pm$ 10° and time-bandwidth products T Δf of about 2000 can be achieved with such dispersive transducers by careful design (88).

For the compression filter in a radar system, the time side-lobe level which would be produced in the compressed pulse by the constant amplitude weighting of Fig. 5.3 would be too great. Thus in the compressor it is common to use some additional weighting over the transducer length, that is over the bandwidth in the frequency domain, which greatly reduces the time side lobes, at the expense of slight pulse broadening, Section 4.2. Figure 5.3c shows a Hamming weighting function chosen to reduce the theoretical time side-lobes below 40 db, and Fig. 5.3e gives the apodizing function for the compression filter including this weighting, the taper to reduce the Fresnel ripples, and the compensation for the roll-off of the uniform transducer and for the frequency scaling.

5.3 Pairs of Dispersive Transducers

The overall bandwidth of a transducer pair such as Fig. 5.2a in which all the dispersion is in one transducer is limited by the bandwidth of the uniform transducer. For wide-band applications this is a serious restraint because the number of fingers in the uniform transducer must be five to ten to have reasonable insertion loss and tolerable spurious bulk-mode generation. This restraint can be removed by using two linearly dispersive transducers as illustrated in Fig. 5.2b (11). Since from

Eqs. 5.8 and 5.9 only a fraction $1/\sqrt{T\ \Delta f}$ of the fingers are effective at a given frequency, the path length traversed and hence the time delay of the different instantaneous frequency components again varies linearly with frequency. If the amplitude weighting of the response is shared equally between the transducers then the frequency scaling factor of Fig. 5.3b will be $(f/f_o)^3$ in each apodization while any Fresnel or Hamming weighting of Fig. 5.3c will be the square of the functions shown. It is very important to note that the overall response of two apodized transducers is not in general the square of the individual apodizing function as noted in Section 4.4, but is so here because the effective fingers at a given frequency "see" only the corresponding effective fingers of the other transducer and both have the same overlap length.

It is also important to note that the concept of one group of fingers in a dispersive transducer interacting only with a corresponding group of a second identical transducer can be used as the basis for the design of wide-band non-dispersive band-pass filters (11). For example, if in Fig. 5.2b one of the linearly dispersive transducers was a translation of the other rather than the mirror image as shown, the path length from effective entry to effective exit would be the same for each frequency and thus the overall frequency response would be non-dispersive. The width of the pass-band is Δf and there are effectively many separate transducers of different frequencies all operating in parallel. The bandwidth is limited by maximum apodization rates, transducer length, second-order effects and manufacturing tolerances rather than by transducer bandwidth of the symmetrical transducers of Chapter IV which

radiate effectively from their centre lines. The steepness of pass-band skirts, the pass-band and out-of-band ripple depend to a large extent on the designer's ingenuity because synthesis techniques for these non-dispersive combinations of dispersive elements are not yet well established.

Many of the second-order effects in the doubly dispersive transducer pairs of Fig. 5.2b result from signal components of a given frequency passing under fingers with which they are not resonant but with which they may interact constructively to produce large reflections or large bulk-wave generation especially serious in wide-band configurations where the bandwidth covers more than an octave of frequency (G4, p.167). As previously noted some of these second-order effects can be reduced through reduced piezoelectric coupling, bulk-wave suppression, multistrip couplers and split fingers. In addition, here where only one group of fingers interacts with the corresponding group in the other transducer, further reduction of second-order effects results from inclining the transducers as indicated in Fig. 5.4. In the case of the broad-band non-dispersive filter the two transducers could be inclined but parallel.

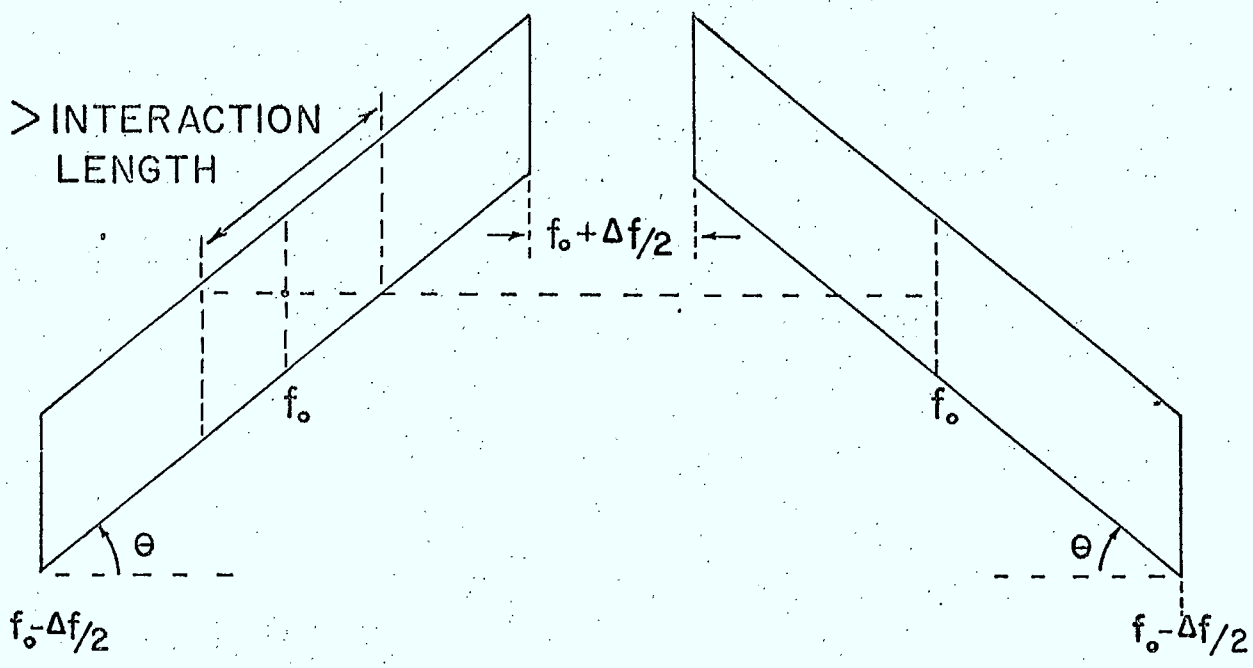


Fig. 5.4 Inclined doubly dispersive chirp compressor.

VI. PHASE-CODED DEVICES

It has been noted already that in radar applications where the peak pulse power at the transmitter is limited, improved signal-to-noise ratio at the receiver can be obtained by using a longer pulse, and that range resolution can be maintained simultaneously through the use of coding within the long pulse and a matched filter and pulse compressor in the receiver. The last section indicated that one form of coding, the linear frequency-modulated pulse, could be both created and compressed by convenient passive surfact-wave devices. Another form of coding which can be used for pulse compression is the phase-coded pulse. Here the outgoing constant-amplitude RF pulse has its phase changed stepwise during the pulse while at the receiver a compression device adds the contributions from all parts of the pulse during some short interval and produces as much cancellation as possible at other times.

6.1 Barker Codes

Compressors for biphase codes are particularly easily realized using surface-wave devices. A commonly used set of codes which produce relatively low, constant amplitude, time side-lobes on a narrow compressed pulse are the Barker codes. For illustration purposes let us consider a received pulse which has been phase shifted according to the five-bit Barker code 10111. This pulse can be transformed to a surface-wave by a broad-band transducer and thus at some instant of time the displacement amplitude would be as shown in Fig.6.1b for the particular case of two RF cycles per bit (35). Now if this wave passes under a transducer with periodic finger spacing corresponding

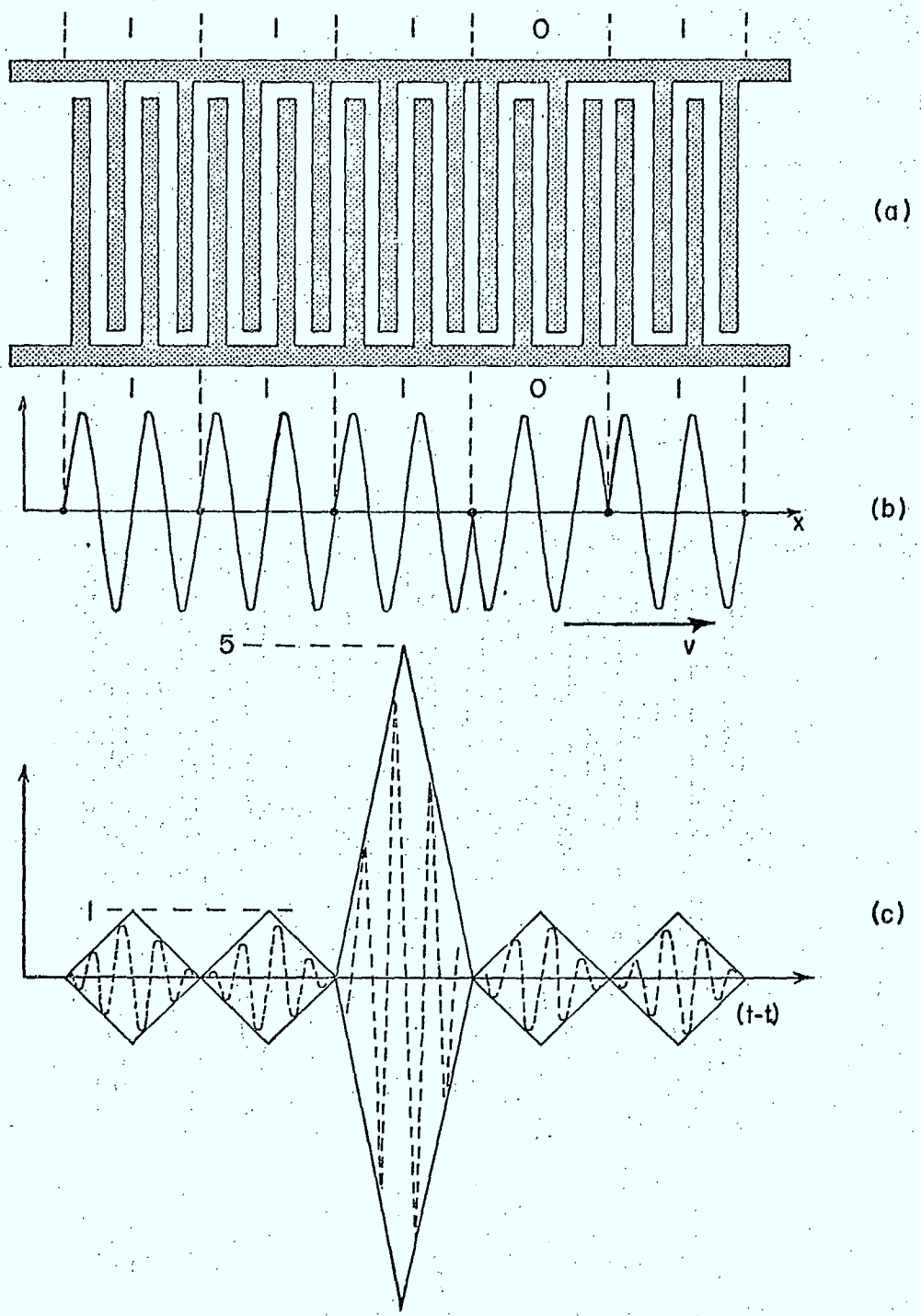


Fig. 6.1 (a) Phase coded transducer for a 5-bit Barker code, 2 cycles per bit.

(b) Surface-wave at instant of correlation t_o.

(c) Electrical output as a function of time in units of $2x/v$.

to the half wave length of this surface wave but phase-coded to the time reversed code 11101 as shown in Fig. 6.1a, then when the wave is precisely within the transducer the voltages generated by every finger pair will add together in phase. The phase reversals in the code are approximated here by connecting adjacent fingers to the same bus, in other types the relative position of a meandering ground bus interleaved between the fingers determines the phase of the different regions (G9). The number of cycles per bit and hence the number of fingers in each periodic region of the transducer is fixed by the system, that is the code length, the operating frequency (RF) and the pulse length.

The electrical output from the transducer as a function of time is illustrated in Fig. 6.1c where the central compressed pulse occurs at time $t = t_o$ when the surface-wave lies completely under the transducer. There is electric output for all of the time that any part of the wave is under the transducer, that is for twice the expanded pulse length. The form of the equal amplitude side lobes is easily verified by conceptually sliding the wave of Fig. 6.1b one bit at a time past the transducer. If the code has M bits, the compressed pulse has a full duration $1/M$ times the expanded pulse length while the side lobes are $1/M$ of the peak amplitude of the envelope as indicated for $M = 5$ in Fig. 6.1c.

Obviously longer codes formed by increasing the number of pulses produce narrower compressed pulses and improve the main-lobe to side-lobe ratio. The maximum number of bits depends on the accuracy and simplicity with which the original pulse can be generated and on the second-order effects in the compressing transducer.

There is no point in increasing the number of bits beyond the value at which second-order effects such as the triple-transit echoes and acoustic regeneration discussed below determine the side-lobe levels and the compressed pulse width.

As was the case with the chirp, the expanded coded pulse for driving the transmitter can if desired be generated by a device of the same general configuration as the compressor. That is, the impulse response of the transducer in Fig. 6.1a is the acoustic waveform of Fig. 6.1b. Thus if an electrical pulse narrower than one half cycle of the fundamental frequency is applied to the transducer, simulating an impulse, the generated acoustic wave of Fig. 6.1b will be radiated in one direction and the complementarily coded pulse in the other, and either of these pulses can be extracted by an appropriate broad-band transducer. Since the amount of electrical energy which can be inserted into the coded transducer during a single impulse is very limited, usually due to interelectrode breakdown, the coded signal amplitude will be small, and such low levels can be a disadvantage for this form of code generation, but this disadvantage may well be outweighted by the frequency and temperature change compensation afforded by the use of similar code generator and pulse compressor (G4, p.270).

The effect of a change in the frequency of the coded pulse is sketched in Fig. 6.2 where the solid curve is the amplitude of the main lobe of the compressed pulse for a 6 cycles per bit, 13 - bit Barker code (35). The relative bandwith between nulls is approximately equal to the reciprocal of the number of fingers, here 1 / 156. It is seen that a one percent frequency shift causes a drastic reduction in

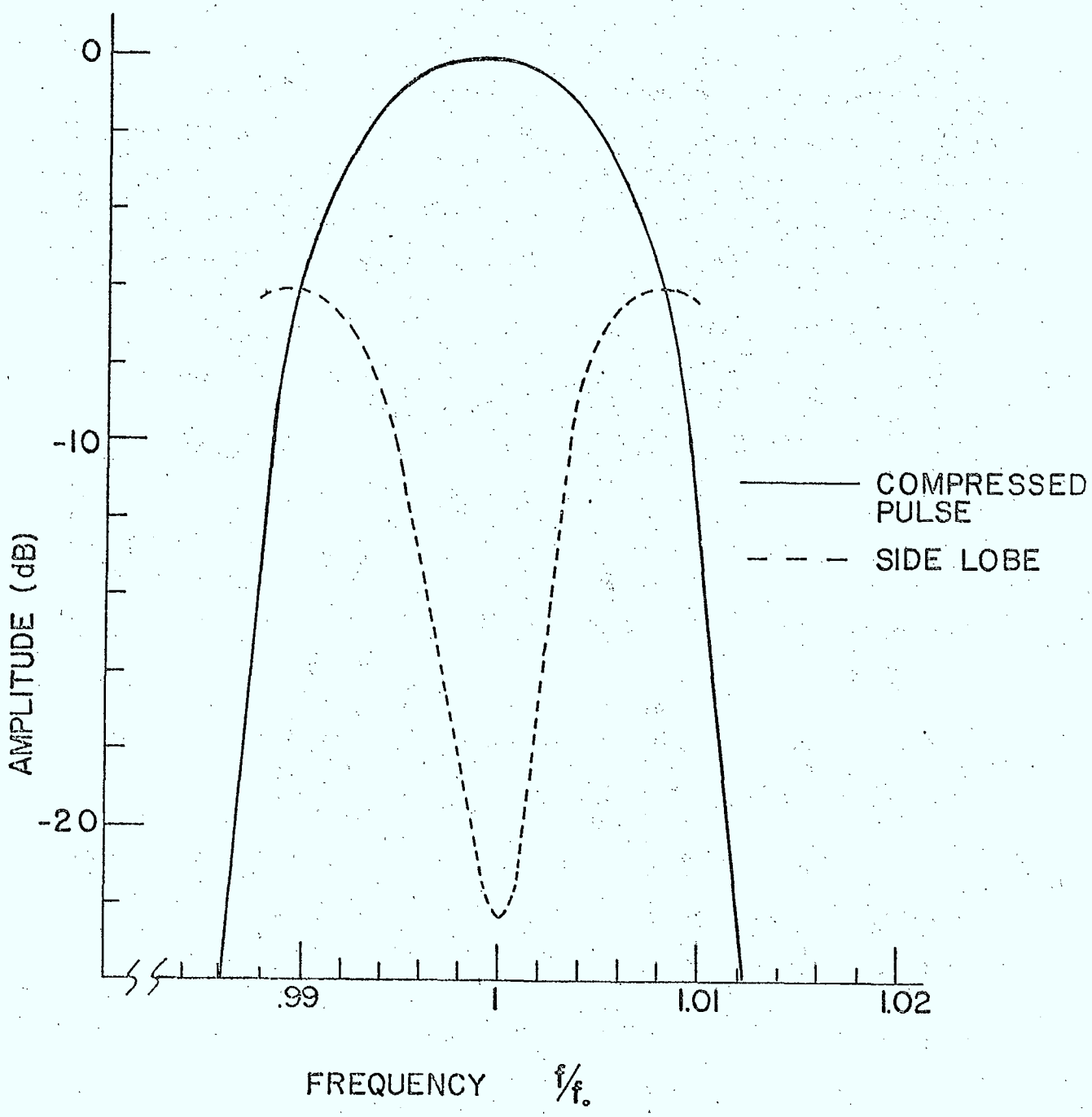


Fig. 6.2 Degradation of amplitude of correlation peak and increase of side-lobes due to frequency shift in a 13-bit Barker code, 6 cycles per bit (35).

the compressed pulse amplitude, but perhaps more important the side-lobe level also rises rapidly because the phase shifts from region to region are no longer the 0 or π required for perfect correlation (10). This question of frequency sensitivity is not confined to the surface-wave type of compressor but is the classic ambiguity function problem in pulse-compression radars subject to doppler shift of the required frequency due to target velocity, and the general problem will not be further discussed. However similar effects would be produced by any shift in the central response frequency of the transducer itself and these are of direct concern here. It is shown in Chapter X that the phase velocity of surface-waves can be temperature dependent, for example the time-delay change of YZ - Li Nb O_3 is 94 ppm/$^{\circ}$C so that a 10° C temperature change would produce a shift of about 0.1% in the central frequency which will produce a significant increase in side-lobe level. Thus to ensure that any effective frequency shifts are introduced by other parts of the system, the surface-wave designer may opt for temperature control of the compressor or for the use of low temperature-coefficient substrate materials; but if the system will allow it, the best technique is to use similar transducers in similar environments or preferably on the same substrate for code generation and for pulse compression, thus producing first-order cancellation of temperature effects.

The metallization to form the transducer electrodes produces reflection of the acoustic wave at each finger edge as discussed in Section 3.3 and here again if the edges are spaced at $\lambda/4$ as in Fig. 6.1 the reflections add in phase and become appreciable. Usually such reflections cause a deterioration in surface-wave device

performance predominantly because of the resulting triple transit echoes. In pulse-compression applications such echoes can frequently be removed by time gating of the output waveform. However the reflected waves under the transducer itself at any instant do not correlate with the transducer and tend to increase the time side-lobe amplitudes, especially the trailing side-lobes, at the expense of the central compressed pulse (36). Again the effects of electrode reflections can be greatly reduced by using split electrodes as in Fig. 3.9 with the increased tolerance requirement in fabrication, or by using a material with a low electromechanical coupling coefficient with its attendant increase in insertion loss.

Another reason for using materials of lower coupling coefficient or for only loosely coupling the transducer to the load is the following: if appreciable energy is extracted from the leading bits of the wave of Fig. 6.1b as it passes under the first regions of the transducer, part of the energy will be reradiated from the transducer as a whole and the remaining wave will be decreased in amplitude with respect to the later bits, both effects produce increased side-lobes relative to the main pulse.

There is an additional degree of flexibility in the surface-wave compressor which can be of use to the device designer. In Fig. 6.1a the number of fingers in each constant-phase region of the transducer was equal to the number of cycles per bit in the coded waveform. This restriction is not necessary as the correlation can be produced with shorter samples of each bit. For example if the third and fourth, seventh and eighth etc. finger were omitted from the transducer of Fig. 6.1b, the

compressed pulse would still have the form of Fig. 6.1c (G10). Similarly if there were six cycles per bit, it would be sufficient to have one properly phased pair of fingers for each bit spaced by six wavelengths. Thus while the number of cycles per bit may be fixed by systems requirements, adjustment of the number of fingers per bit allows some compromises between bandwidth, insertion loss and reflections. Incidentally if the number of finger pairs per bit is less than the number of cycles per bit and the transducer is used as the code generator the driving electrical "impulse" will have to contain the appropriate number of cycles to fill the bit interval in the wave on the substrate.

The discussion here has been illustrated by Barker codes which are attractive for pulse compression because of the relatively low, equal-amplitude, side-lobes. However, apparently there exists only a limited number of such codes (maximum 13 bits) and for larger pulse-compression ratios it is necessary to go to other forms of sequences which have similar properties but not such simple compressed waveforms as the Barker, Fig. 6.1c (10, 15). However, any biphase coded sequence can be implemented by the surface-wave transducer using the phase reversal techniques of this section, a statement which leads us to another communications application as discussed in the next section.

6.2 Spread Spectrum

There is a type of pulse communication system of limited but important application. In it the individual data bit is encoded into a very broad band pulse stream, the encoding is chosen so that the transmitted signal resembles random noise.

For example if a data bit "1" is transmitted as a binary-phase coded sequence, say a 127 - chip pseudo-noise sequence, then this transmitted signal has a wide flat frequency spectrum over a bandwidth of the order of twice the chip rate, the rate at which the carrier phase is switched, about the carrier frequency (G4, p. 309 and 319). The data bit "0" could be transmitted by the complementary code. Thus on the transmission link the signal resembles noise over a bandwidth very wide relative to the data rate. At the receiver however, if the transmitted code is known, a 127-chip correlator of the type shown in Fig. 6.1a will reproduce a compressed pulse for each transmitted sequence.

The obvious disadvantage of such a spread-spectrum communication system is the extremely wide bandwidth required. However the processing gain can be large so that many such channels operating over the same frequency band with different codes produce relatively small interference, or one spread-spectrum channel might be added to the composite signal of some other multichannel link where it would but slightly increase the noise level of all the other conventionally modulated channels. Such spread-spectrum signals are very difficult to jam because they are relatively insensitive to narrow-band signals. Also they offer a degree of privacy against unauthorized use because of the need for the receiver to know the transmitted code to receive the message and a degree of low detectability because the signal is often below receiver noise levels. For very secure operation, long and changeable codes are required and experimental surface-wave devices for the latter case are considered later.

The generation of the transmitted code follows the principles discussed in connection with the phase-coded pulse compressors for radar (5, 30). For example if the transducer of Fig. 6.1a were phase coded according to a long pseudo-noise sequence instead of the short Barker code illustrated there, an applied impulse shorter than a cycle of the centre frequency would produce an acoustic wave in one direction phase-coded according to the pseudo-noise sequence, perhaps two cycles per constant phase region or chip of a 127 - chip code. A wave having the complementary code is radiated in the opposite direction and thus if two broad-band transducers are placed equal distances from each end of the coded transducer, then for each impulse a burst of the pseudo-noise sequence and its complement are available simultaneously at the transducer outputs. One code can be transmitted for the data bit "1" and the other for "0", and if the impulse repetition period is exactly equal to the sequence length and the data switching is similarly synchronized there will be no appreciable periodic component in the transmitted signal and its spectrum will resemble that of broad-band noise.

The received signal can be converted to an acoustic wave by a broad-band transducer and passed under a transducer phase coded for the pseudo-noise sequence used, and each time a data bit "1" is transmitted a compressed pulse is produced at the coded transducer. The same received signal propagated in the opposite direction under a similar transducer will produce a compressed pulse for each complementary sequence, that is for each data bit "0". Envelope detection and subtraction of the two transducer outputs reproduces the original data train (G4, p.309). Techniques

exist which allow a single surface-wave device to be used to produce distinguishable correlation peaks for the "1" and "0" data bits.

It is important to realize that with the surface-wave correlator in the receiver, a compressed output pulse results from the received coded pulse regardless of the timing of the latter. This feature is not possessed by many other forms of correlators which involve switching, and is an asset in acquisition time and synchronization for spread spectrum systems.

Since the desirable codes for spread spectrum applications are long, the corresponding transducers are long, and if low insertion loss is important the individual fingers must have high electrical coupling to the substrate with the result that colinear transducing regions of the form of Fig. 6.1 can lead to large second-order effects due to signals generated in one region passing under others. Such problems would be greatly simplified if each constant phase region were displaced laterally from its predecessors as shown in Fig. 6.3, all fed from a common electrical input and all feeding in turn a wide output transducer. Each block represents the number of fingers appropriate to the given number of cycles per chip and with bus connections to give the appropriate phase. A more practical arrangement to produce essentially the same parallel track arrangement but with track overlap is shown by the broken lines of Fig. 6.3. Here again complementary codes are generated in opposite directions.

Phase-code generators and correlators of the inclined type are available with reproducible and system-useable characteristics in the 200 Mhz range with 127 chips, 5 cycles per chip (37), while experimental units up to 511 chips have been fabricated.

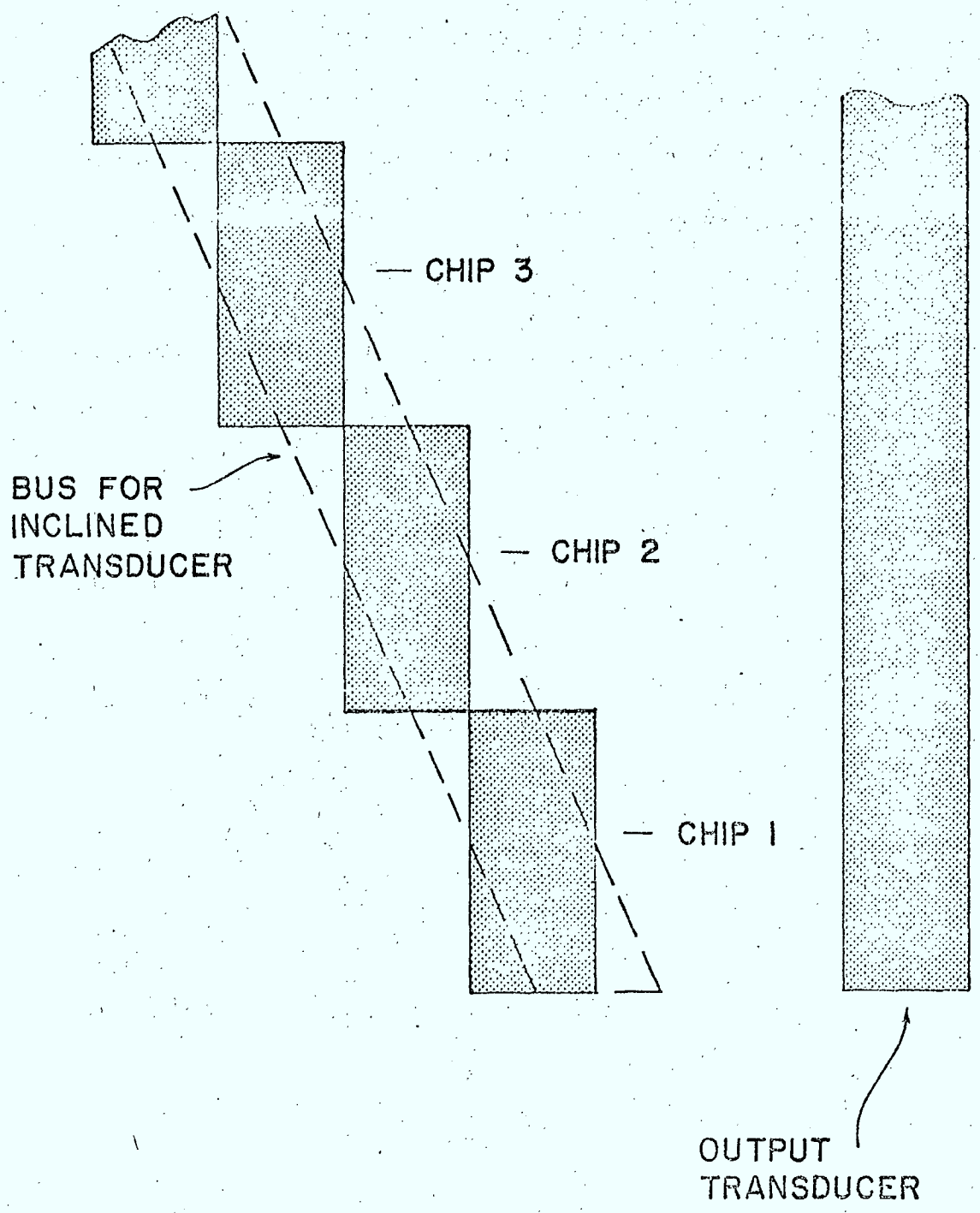


Fig. 6.3 Staggered or inclined phase-coded transducer.

6.3 Programmable Phase-coded Devices

It is obvious that the fixed codes discussed above are not long enough for security in the case of intelligent jammings where a jammer attempts to cancel the processing gain advantage by generating jamming signals with characteristics similar to the wanted signal. But even in the latter, surface-wave devices offer attractive possibilities when used in conjunction with some form of electronic switching. For example a surface-wave linear FM compression filter (Fig.5.3) could be used as the matched filter in a frequency-hopped chirp system (G6, p.494) or a surface-wave convolver as discussed in Chapter XIII could be used to provide an ever-changing reference waveform. But here under phase-coding, the programmable phase-coded delay lines should be discussed. In Figure 6.1a it was seen that the "phase" of a given section of a transducer depended on the relative connections of the fingers to the two bus bars in different regions. Obviously if in a given region the connected ends of the fingers could be disconnected at the bus bars and the free ends connected to the neighbouring bus the phase of the region would be reversed.

The form of programmable tapped delay line or code generator is shown in Fig.6.4. Again the electrical signal is applied to or extracted from the two bus bars but here the bus to which each finger is connected is set by the polarity of the bias voltage applied to that finger thus in operation the phase, and conceptually the electrical weighting, of each finger can be electronically switched. Such an arrangement gives an immense versatility to the choice of phase codes available. The changing of finger weighting by changing electrically the effective resistance between the fingers separately and the bus bars has not been exploited as yet, though MOSFET

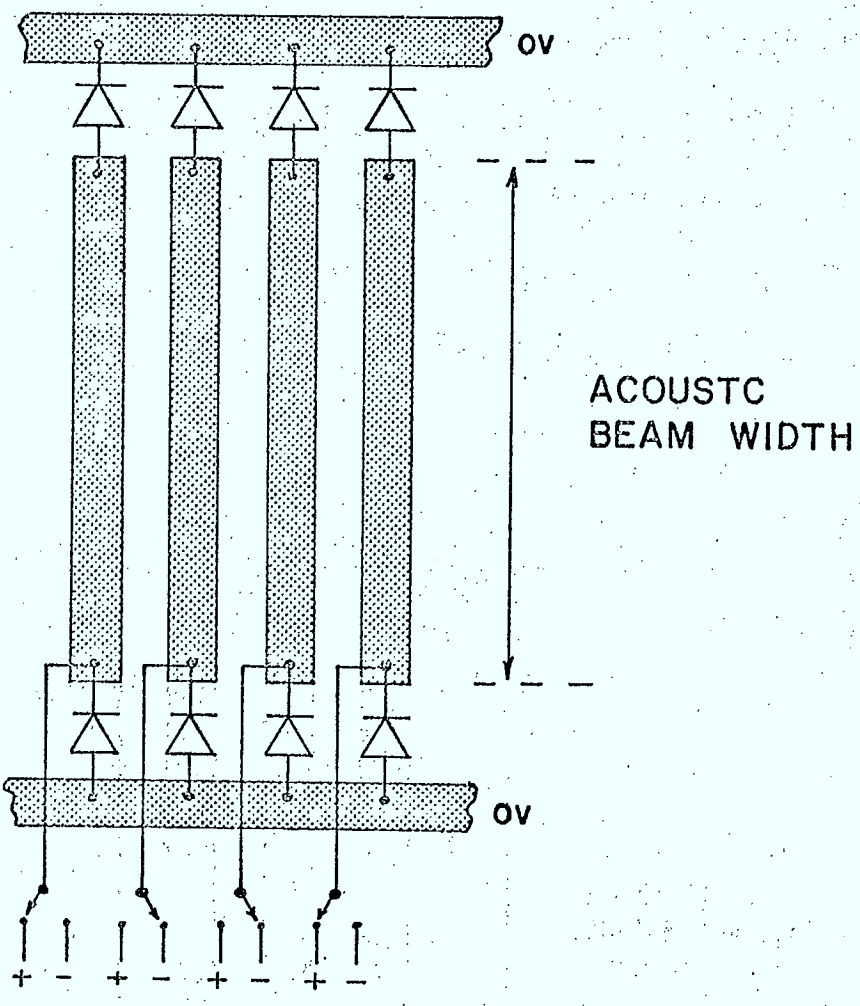


Fig. 6.4 Programmable phase-coded delay line.

geometry has been used on silicon substrates to provide voltage-controlled piezo-resistive taps (12).

The on-off character of diodes driven from cut-off to conduction as is required for phase coding works well with transducers of the form shown in Fig. 6.4 using discrete diodes (31, 69, 91). However the use of discrete diodes wire bonded to the fingers and separate switching sources is not very practical and the device designer must turn to integrated techniques.

The most promising method for integrating surface-wave transducer and semiconductor devices seems to lie with aluminum nitride films on sapphire substrates. The Al N films are piezoelectric and these films on sapphire form an acoustic medium of reasonable propagation characteristics (Section 10.5), and most important, silicon films can be grown on the same substrate (G5, p.274). Thus the transducers can be deposited on the regions of the surface covered by an AlN film while the diodes and transistors are fabricated in the adjacent silicon films. Experimental versions of such programmable integrated surface-wave delay lines have been demonstrated with 64 taps but detailed performance characteristics are not available (G4, p.92), (G6, p.333).

VII. REFLECTIVE ARRAYS

There is another deceptively simple concept which has been carried over from bulk-wave to surface-wave dispersive delay lines and which offers distinct advantages over weighted transducers for large time-bandwidth devices. The concept is to use the resonant reflection from a periodic perturbation to define the length of the transmission path, thus controlling the variation of time delay with frequency.

The perturbations of the surface can be produced in several ways such as overlaid deposited strips (see Section 3.3) but the method which produces the smallest second-order effects uses shallow grooves etched directly into the substrate surface. The prototype geometry of a reflective-array compressor (RAC) is shown in Fig.7.1 (106), (G4, p.181), (G6, p.490). Here the input and output transducers are assumed for the moment to be very broad band. The spacing of the etched grooves varies linearly with distance along the substrate with the result that the different frequency components suffer the double reflection at different regions of the reflector and thus the path length transversed by a signal component is a linear function of its frequency. Before commenting on the advantages of this method of pulse compression let us first examine some of the design parameters.

7.1 Design Considerations

For a single step of height h on a substrate the amplitude reflection coefficient r of a normally incident surface-wave is proportional to the step height for steps which are very small with respect to the wavelength. For YZ $LiNbO_3$

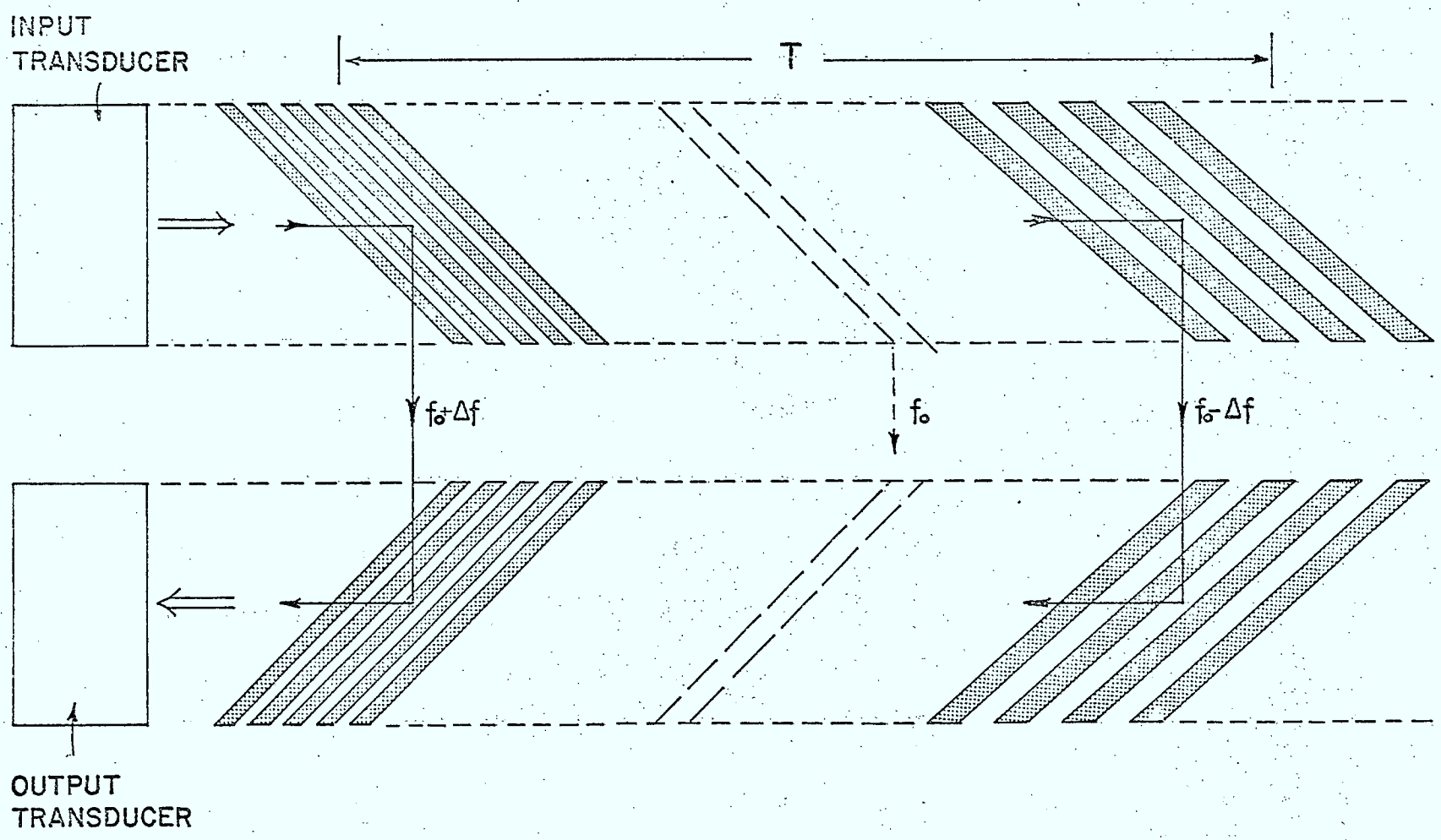


Fig. 7.1 Geometry of a reflective array compressor. Inclined grooves are etched into substrate surface.

$$r \simeq \frac{1}{3} h/\lambda \qquad h/\lambda < 10^{-2} \qquad (7.1)$$

and this value can be used for the reflection from the angled up-steps and down-steps of the reflective array (G5, p.263). It is desirable that the time taken for a frequency component to traverse its path be dependent only on its point of reflection along the substrate and not on the number of other grooves it has traversed, that is the grooved path should be effectively non-dispersive. It is found that for YZ - Li $Nb O_3$ the dispersion, change in velocity of propagation with frequency, is negligibly small providing the groove depth remains less than $10^{-2}\lambda$.

In the discussion on dispersive interdigital transducers it was found (Eq.5.9) that at any given frequency a certain effective number of fingers were radiating. Similarly here, for a linear FM spacing the effective number of grooves is (G4, p.183)

$$N_e = f\sqrt{T/2\Delta f} \qquad (7.2)$$

again using a band of grooves within which the phase shift is less than $\pi/2$ (see Eq.5.9) as a criterion for effectiveness. The frequency response of the RAC is thus given to a very good approximation by

$$|H(f)| = g(f)\,[N_e\, r(f)]^2 \qquad (7.3)$$

where g is a geometric factor dependent on the effective width of the beam for one frequency component as it traverses between the grooved tracks ($N_e\lambda$) relative to the beam width W on the tracks; $g = \frac{W}{N_e\lambda}$ for $\frac{W}{N_e\lambda} << 1$ and in the other limit $g = 1$ for $\frac{W}{N_e\lambda} >> 1$.

One novel and important characteristic of these double track reflective-array devices is that the different frequency components are spatially separated during part of their travel. Thus frequency weighting could be introduced along the centre line between the tracks. This method has not been used for amplitude weighting but has been used for phase correction of fabricated devices by plating a varying width strip in the central region such that the delays in the transverse direction are perturbed, due to the decreased phase velocity, by just enough to compensate for the previously measured phase error at each frequency (G6, p.490).

Amplitude weighting is usually accomplished by varying either the depth of the individual grooves or the groove length across the acoustic beam, or by varying both. For relatively gradual weighting of amplitude with frequency, as met in pulse compressors, varying the groove depth is the most suitable technique. While it might be advantageous from the point of view of insertion loss to have relatively deep grooves, the practical limit is the $10^{-2}\lambda$ mentioned above, otherwise the grooved tracks become a dispersive medium. Thus during fabrication the groove depth can be varied in the direction of propagation, by controlling the etching time, from effectively zero depth to the above limit. The desired overall weighting functions depend on the application and for pulse compression might include the Hamming weighting for side-lobes illustrated in Fig. 5.3c. For reflective arrays, just as for dispersive transducers, Fig. 5.3b, there are inherent frequency dependencies which should be compensated in the groove weighting. Here if the tracks are wide in terms of the wavelength there is an f^4 factor in Eq. 7.3 which must be included in the

design along with compensation for the roll-off of the input and output transducers (G4, p.183).

On an isotropic substrate the angle of the grooves with respect to the track direction could be 45°. The substrates for high frequency devices are single crystals and thus not elastically isotropic. If the velocity of propagation is not the same in the track direction as perpendicular to it, the angle of the grooves must be adjusted from 45° to produce the total 180° - reflection (106). This condition requires a very precise knowledge of the surface-wave velocities to ensure that the input and output acoustic beams are parallel and also precise crystallographic orientation to ensure device reproducibility with a given photo mask (G5, p.323).

The freedom from spurious signal interference of the reflective array comes dominantly from its geometry. The very resonant reflection which causes so much difficulty with standard transducer pairs, here forms the operating mechanism of the desired signal itself. Recalling that the grooves are but perturbations, it is seen that to get from input transducer to output transducer, other than by electrical feed through, the signal energy must satisfy Bragg reflection conditions twice in such a manner as to have a total reflection angle of 180°. Second-order paths meeting these conditions are very improbable. Moreover, on an anisotropic substrate where the groove angles are set to match Bragg conditions for the ratio of the two perpendicular surface-wave velocities, a bulk wave will have a different ratio of velocities and thus will not be able to find a path at any frequency which satisfies its Bragg conditions twice and returns parallel to the second track. Finally, the reflective arrays do not suffer from the reradiation phenomenon of dispersive interdigital transducers.

The delay time of a surface-wave depends on the temperature for most substrates (65). In linear FM dispersive lines as under discussion here, such changes in temperature change the slope of the time-delay vs frequency curve because the path length varies linearly with frequency. Such slope changes are not important if similar devices at the same temperature are used for both pulse expansion and compression. However if the expanded chirp is electronically generated, the substrate temperature coefficient of the compressor must be small or the temperature controlled to maintain a slope matched to the expandor.

Since the operative mechanism of the reflective array devices is a mechanical perturbation, the part of the substrate containing the grooves need not be piezoelectric and thus this part of the substrate material can be selected for properties other than electromechanical coupling coefficient, for example it may be chosen for small coefficient of time delay with temperature. Even if the reflective region is non-piezoelectric the transduction region must be. Efficient surface-wave generation has been achieved both in bonding Li Nb 0_3 to the ends of the non-piezoelectric part of the substrate, polishing the common surface and depositing the interdigital transducers on the Li Nb 0_3 region (G4, p.11), (G5, p.365 and 367), and in depositing piezoelectric overlays onto the substrate onto or under the transducer (28) see Section 10.5.

7.2 Examples

Four examples of realized reflective-array compressors are shown in Table 7.1 (G4, p.185). The first two compressors are up-chirps while the latter

TABLE 7.1

REFLECTIVE ARRAY COMPRESSORS (G.4, p.185)

Time Bandwidth $\Delta f \cdot T$	Bandwidth Δf MHz	Dispersion T μ sec	Centre Frequency M Hz	R M S Phase Error (degrees)	Matched C W Insertion Loss (dB)
120	6	20	60	0.87	26
600	6	100	60	1.5	34
1500	50	30	200	3.5	38
5120	512	10	1000	7.4	~50

two are down chirps. The most striking features of these devices are the extremely small departures of the phase vs frequency curve from true quadratic shape over the thousands of wavelengths involved. In the last device, careful control of fabrication gave an R M S phase error of $\pm$ 18° which was compensated afterwards by a plated variable width metal strip along the device centre line to the error shown. Other similar devices have had R M S errors less than $\pm 3^{\circ}$ (G6, p.492). One indication of the low level of spurious signals is that the side-lobe levels of the first three devices were very close to the theoretical 35 dB appropriate to the weighting function, and even the device with $\Delta f\, T$ of over 5000 had experimental side-lobes of about 25 dB One serious disadvantage of the reflective arrays is the large insertion loss shown in Table 7.1 .

Figure 7.2 (G4, p.189) indicates the expected limits of surface-wave reflective-array compressors. The dispersion limit is set by the available size of homogeneous substrates while the maximum bandwidth is set by both line width tolerance in the etching and groove position accuracy. Second-order effects such as diffraction and inhomogeneities in velocity indicate the time-bandwidth limitation of 10^4.

7.3 Reflective Array Filters

There are other forms of the reflective arrays which while they have not been extensively exploited as yet are of potential usefulness. Consider for example the geometry of Fig. 7.3 which is the reflective array equivalent of the non-dispersive band-pass filter of Section 5.3. Here each frequency component traverses the same path length in going from input to output transducer and thus the overall transfer

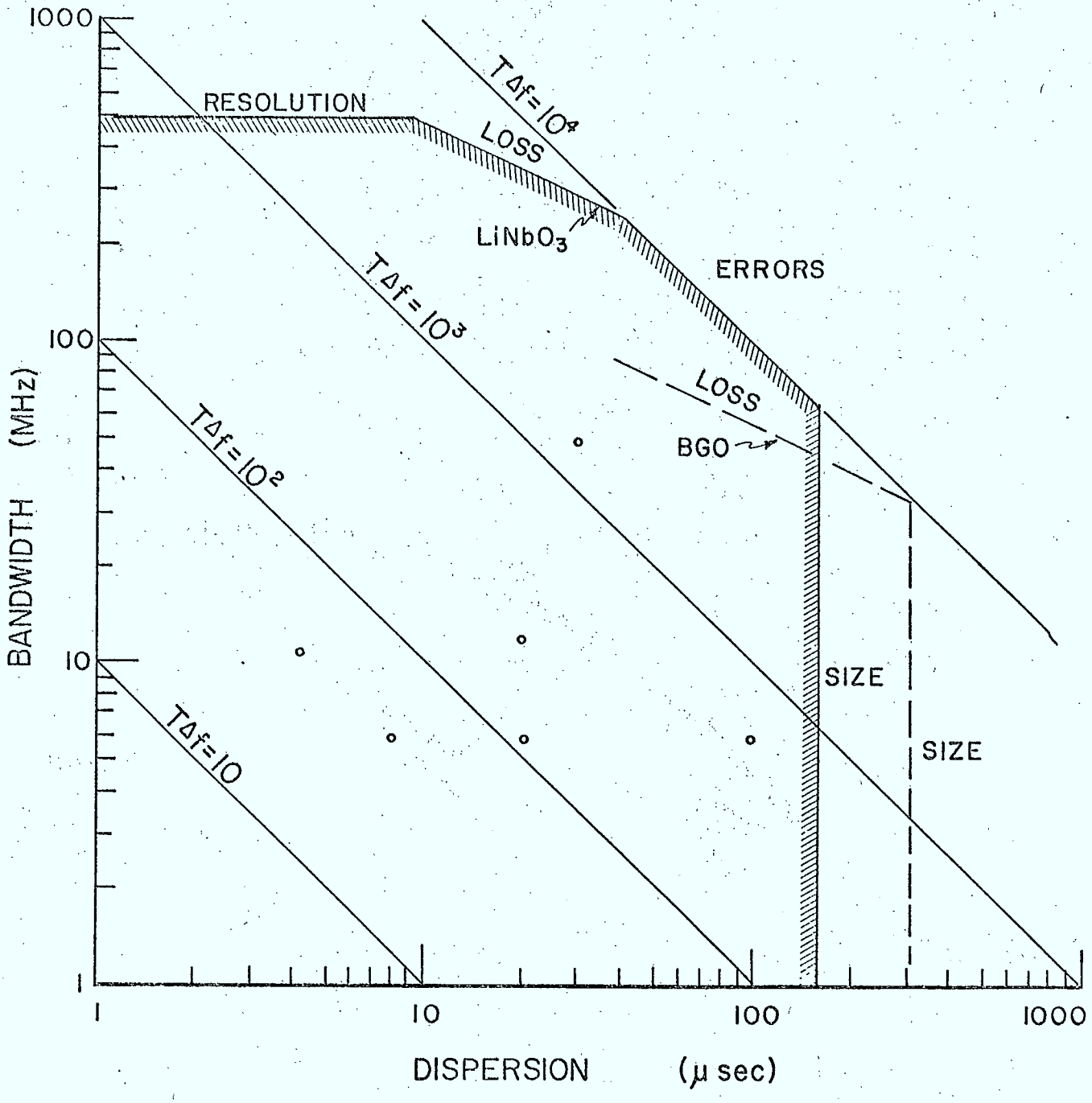


Fig. 7.2 Expected performance limits of reflective array compressors. Circles indicate operating devices.

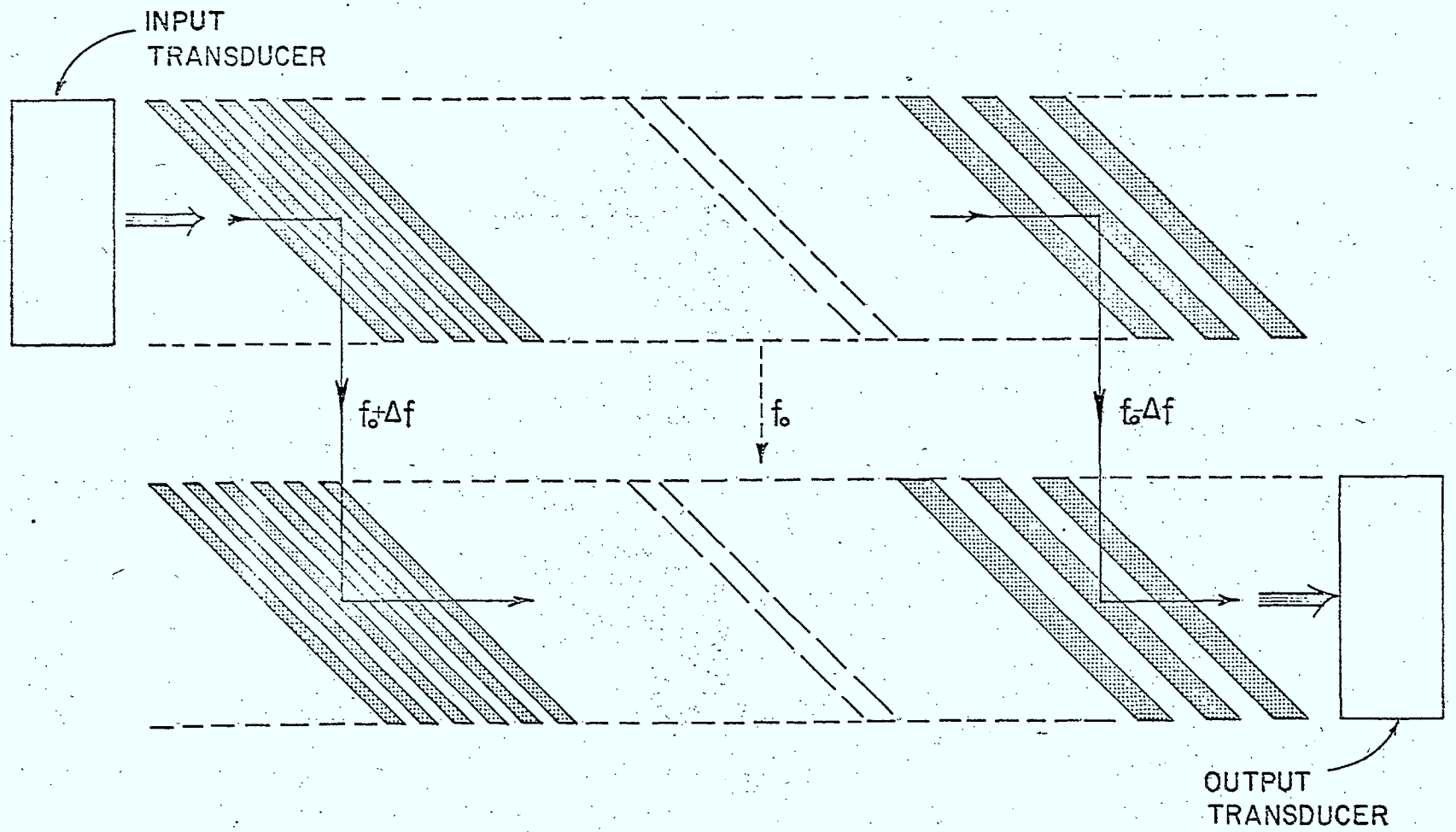


Fig. 7.3 A reflective array in a non-dispersive band-pass filter configuration.

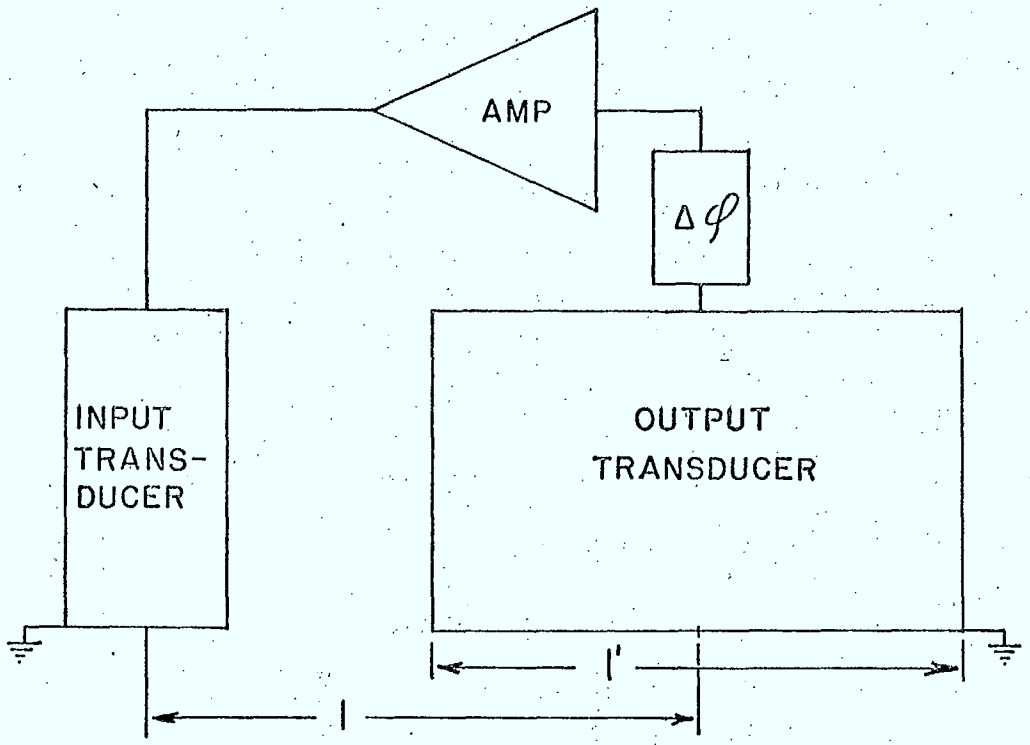
characteristic is non-dispersive. The amplitude weighting of each frequency component in the pass-band will depend on the reflection coefficients of the grooves which are effective for that frequency. The weighting can be done by varying either the length or depth of the grooves or both. Such arrays should be useful in wide-band filters if the insertion loss can be tolerated.

It should also be noted in Fig. 7.3 that if the output transducer is on the same track as the input, we have a band-stop configuration complementary to that obtained with the second-track as the output. In particular if the grooves are almost equally spaced and if the output is taken on the same track and the signal on the second track is absorbed, a notch-type rejection filter is produced.

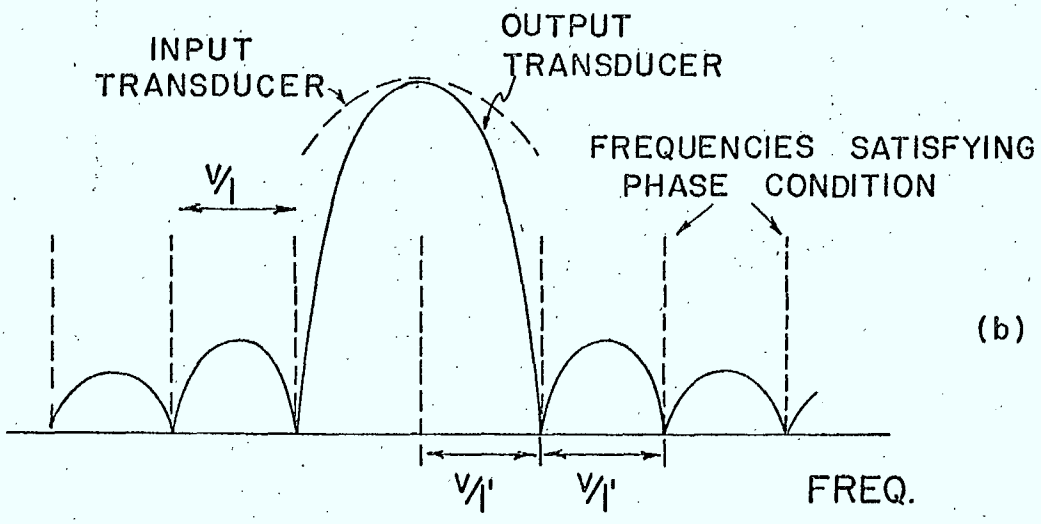
VIII. OSCILLATORS

The use of a surface-wave delay line as the frequency determining element in an oscillator offers many attractive features which are now beginning to be exploited (14), (G4, p.63). If constructed on a low-temperature coefficient cut of quartz the oscillator will have stability comparable to that of a conventional quartz crystal-controlled oscillator but can operate with a fundamental frequency orders of magnitude higher and can provide a much more rugged unit of smaller size which is compatible with planar technology and integrated active circuits. Moreover such an oscillator can be frequency modulated, or at least electronically tuned, and the output can be tapped conveniently at any desired phase.

The principle of the oscillator is illustrated in Fig.8.1. If the two transducers are symmetric then each effectively radiates from its centre line, Section 4.3, thus with a spacing of l between these centre lines and an amplifier gain sufficient to overcome the insertion loss, the oscillation condition is satisfied at each of the frequencies spaced at v/l; Fig.8.1b. The absolute values of these frequencies will depend on the phase shift represented by $\Delta\phi$ in the electronic part of the feedback loop. Now the nulls of a uniform transducer of length l' are spaced at v/l' (Eq.3.6) in frequency. Thus if $l' = l$ and the incremental phase is set so that one of the frequencies satisfying the oscillation condition say f_{os} is at the centre frequency of the long output transducer then all the other frequencies satisfying the phase condition for oscillation are at the nulls of the transducer response and thus



(a)



(b)

Fig 8.1 (a) Surface-wave oscillator configuration.

(b) Frequency response of the output transducer.

cannot have sufficient loop gain for oscillation. The broader band input transducer might also be used to further reduce the possibility of oscillation at spurious frequencies by judicious location of its nulls on the side-lobes of the longer transducer. There are of course many choices other than uniform for the configuration of the pair of symmetrical transducers which will have one central lobe in the overall response and nulls equally spaced at v/l.

Providing the gain is not too high, the oscillator frequency can be shifted approximately $\pm v/4l$ by variation of the electronic phase shift by angles $\pm\pi/2$. Such changes vary the loop gain because of the shape of the transducer response, but if changes in loop gain are important they can be reduced by making the top of the transducer pass-band flatter through apodization.

On an ST quartz substrate, the substrate with the lowest temperature coefficient of delay currently available, the frequency shift of a surface-wave oscillator is (G4, p.63)

$$\frac{\Delta f}{f_o} = -3 \times 10^{-8} (\Delta T)^2 \qquad (8.1)$$

over a wide range of temperatures, where f_o is the design centre frequency at the design temperature, $\Delta T = 0$. Thus the temperature stability is comparable to that of a quartz crystal oscillator, and as in the latter case, temperature control is needed if stabilities of the order of one part per million are required.

As with most other surface-wave devices, surface-wave oscillators are most conveniently used in the frequency range 30 Mhz to 2 GHz, limited at the lower

end by substrate size and at the upper end by fabrication tolerances for fundamental operation and by losses for harmonic operation.

Fully integrated surface-wave oscillators have not yet appeared but they can be anticipated shortly. If it were not for the large temperature coefficient involved, it would appear that Al N layers on sapphire would be attractive substrates for such integration of the active and passive components of the oscillator.

Because of the relatively long time delay between the transducers, the surface-wave oscillator cannot be modulated very fast, maximum modulating frequency is of the order of $v/10l$ (G6, p.344). However, outputs of any desired phase, say two quadrature components, can be extracted by appropriately spaced taps on the acoustic path either between the main transducers or just beyond the input or output. Again various techniques can be used to minimize spurious responses.

IX. MULTISTRIP COUPLERS

There is a surface-wave concept which is of broad potential utility in devices but suffers greatly in its current applicability to only high coupling substrates. In several previous sections the metal electrodes of an interdigital transducer have been considered as sampling elements which sample a transverse section of the acoustic wave passing beneath, and by means of electrical connections transfer these samples to an electrical signal in a load. However, these individual samples could be electrically connected to other radiating fingers on the piezoelectric substrate and thus extract energy from one acoustic beam and re-radiate it into another.

9.1 Principle of Operation

In the prototype configuration, the two beams or tracks are side-by-side and the individual fingers over one track continue over the other as shown in Fig. 9.1 and thus the two sets are electrically connected at their centres (57, 58, 59, 60). In the region under the strips in the prototype we have a geometry which is symmetrical about the centre-line between the two tracks, thus the modes of propagation can be separated into symmetric and antisymmetric modes.

If for purposes of discussion it is assumed that the input wave on track 1 is uniform across the track, only system eigenmodes which are also uniform across track 1 at the input need be considered. Since the strips are assumed to have no resistance, each must be equipotential along its length, there are thus two eigenmodes of propagation. For the symmetrical mode the particle displacements are in phase on each

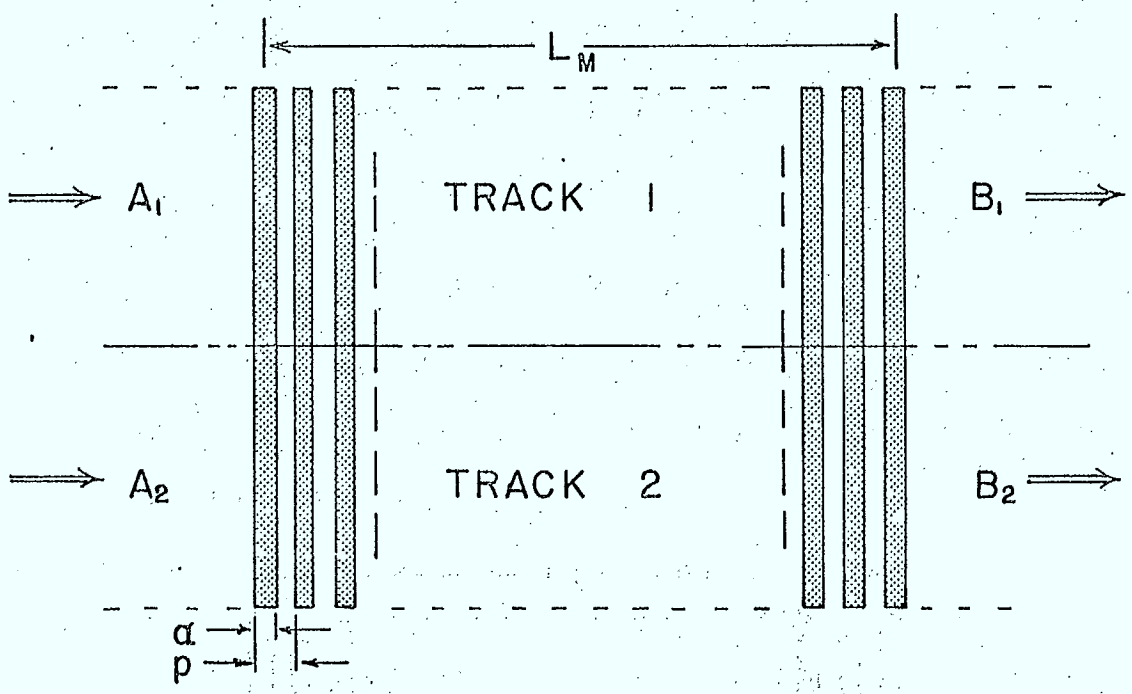


Fig. 9.1 Two acoustic tracks coupled by a multistrip coupler.

track and the electric potential of a strip oscillates in synchronism with these displacements. If the strips were of zero width in the direction of propagation their presence would have no effect on this wave which is uniform across the two tracks and its velocity would be that of a surface-wave in the same direction on a free surface. However the finite width of each strip forces the potential at each instant to be constant in space over each strip width rather than varying exactly sinusoidally in space as on a free surface, with the result that the velocity of this symmetrical mode is somewhat less than the free-surface velocity.

For the antisymmetrical mode the particle displacements are 180° out of phase on the two tracks so the potential of each strip remains always at zero though there are currents back and forth along the strips as the wave progresses. If the strips were as wide as their period, this mode would have the same velocity as that of a surface-wave in the same direction on a shorted surface. However the strip width is less than this maximum and the surface potential between strips need not remain zero, with the result that the antisymmetric mode has a velocity somewhat greater than the shorted surface case.

Any input wave which is uniform across track 1 and uniform across track 2 at the input to the coupler can be decomposed into a sum of the two normal modes above, and these modes propagate independently and unaltered through the region of the strips, and at the output can be added on track 1 and on track 2 to provide the respective outputs. Since the phase velocities of the two modes are different, the output amplitude on track 1 will differ from the input amplitude on the same track,

similarly for track 2. The phase difference between the antisymmetric and the symmetric modes at the output will be (6, 7, 56)

$$\varphi_L = (k_A - k_S) L_M = \frac{R_k}{2} K^2 k_f L_M \tag{9.1}$$

where k_f is the wave number on a free surface and R_K is a geometric factor for finite strip width sketched in Fig. 9.2. So that if A_1 and A_2 represent complex amplitudes of some component of the input wave on tracks 1 and 2 respectively while B_1 and B_2 represent the output amplitudes, it is easy to show that the matrix relation between them is (G6, p.155)

$$\begin{bmatrix} B_1 \\ \\ B_2 \end{bmatrix} = \begin{bmatrix} \cos(\varphi_L/2) & -j\sin(\varphi_L/2) \\ \\ -j\sin(\varphi_L/2) & \cos(\varphi_L/2) \end{bmatrix} \begin{bmatrix} A_1 \\ \\ A_2 \end{bmatrix} \tag{9.2}$$

where the absolute phase shift appropriate to the average phase velocity of the modes has been omitted.

Now if there is an input signal on track 1 only, $A_2 = 0$, and if L_M is chosen to produce a phase shift $\varphi_L = \pi$ between the symmetric and antisymmetric modes, then $B_1 = 0$ and $|B_2| = |A_1|$ so that all of the incident energy on track 1 is transferred to track 2. On the other hand, if again $A_2 = 0$, but the coupler length is chosen for a $\pi/2$ phase difference between modes the outputs are $B_1 = A_1/\sqrt{2}$ and $B_2 = -jA_1/\sqrt{2}$, neglecting the common phase shift, and thus we have a 3 - dB power divider.

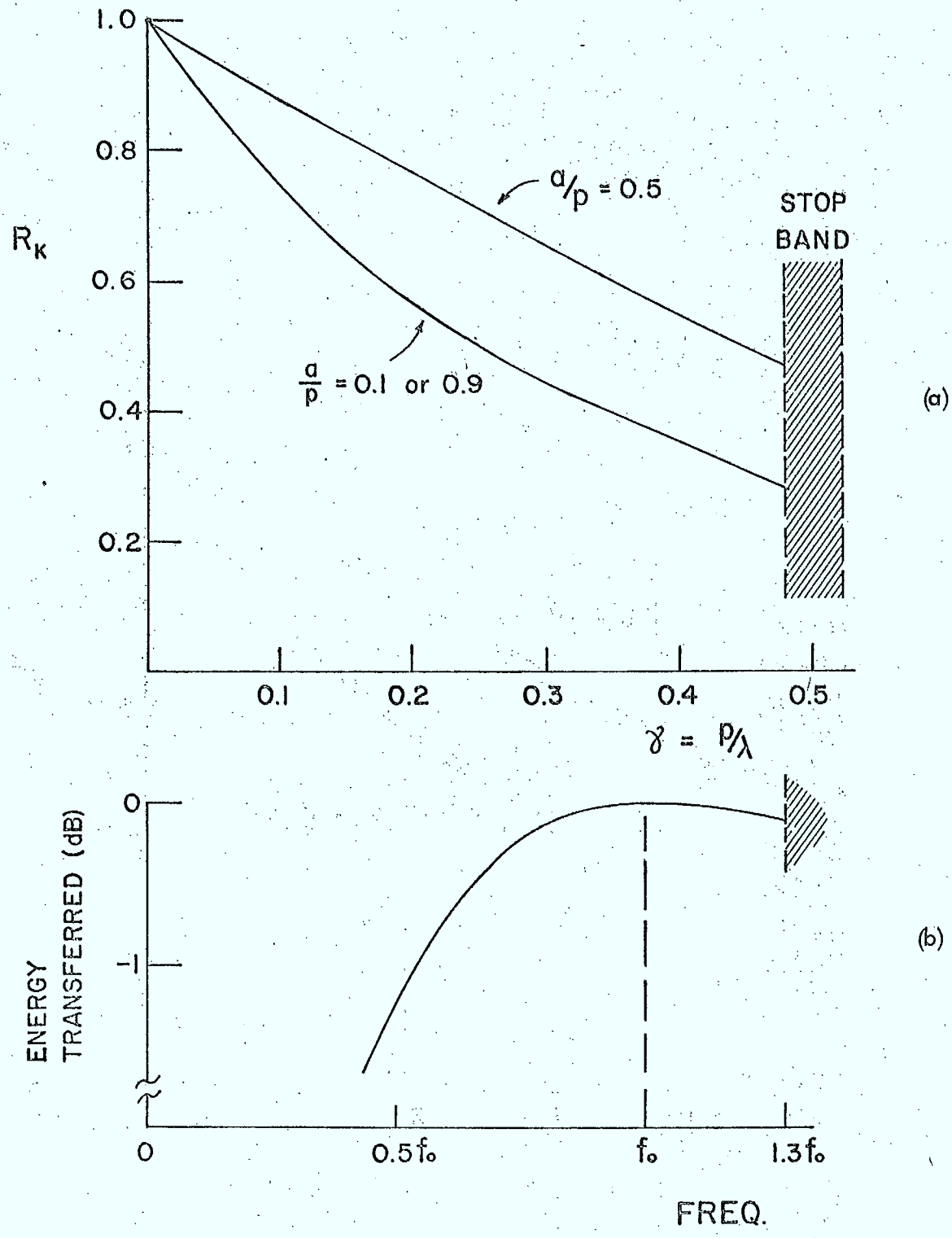


Fig. 9.2 (a) Coupling factor R_K in Eq. 9.1 as a function of the number of periods per wavelength.

(b) Theoretical frequency response of a track changer designed for complete transfer at $\gamma = 0.375$.

These roles of track changer and of power divider form the basis of a range of possible surface-wave components a few of which are mentioned below, but first let us consider the Equations 9.1 and 9.2 more quantitatively.

9.2 Track Changer

The abscissa in Fig. 9.2 is the reciprocal of the number of strips per wavelength and thus at a given frequency it is desirable from a fabrication tolerance standpoint to have p/λ as large as possible even though the thereby reduced R_K will require more strips. The optimum strip width "a" is one half the period as indicated by the curves for three different metallization ratios in Fig. 9.2. For p/λ close to 0.5 (two samples per wavelength), the backward wave associated with the space harmonics of the periodic structure becomes important and there is a "stop band" within which the structure does not propagate, see Section 7.1 (59, 60). The structure propagates again for p/λ above 0.5 but in this range there is a large coupling to bulk modes and the coupler is unlikely to be useful in systems.

As an example consider a multistrip coupler on YZ - Li Nb 0_3 chosen to have $\gamma = 3/8$ and $a/p = 0.5$. The value of R_K from Fig. 9.2 is 0.58 so that from Eq. 9.1 the length for a phase shift $\varphi_L = \pi$ is

$$L_{M\pi} = \lambda / R_K K^2 \cong 38\,\lambda \qquad (9.3)$$

that is 102 strips. This number of strips is practical on Li Nb 0_3 especially because a few broken or shorted fingers do not seriously degrade the performance. However, the number of strips required becomes excessively large for materials with low coupling

constants, for example this same track changer on ST quartz would require some 4000 strips.

When it is useful and useable, one of the advantages of a multistrip coupler is the relatively broad band width of the device. The above track changer would have the frequency response shown in Fig. 9.2b. Here the response at the high frequency end is set by the stop band. The response could be made more symmetrical about the centre frequency f_o by choosing $\gamma = 0.25$ (127 strips) but the four strips per wavelength might be difficult to fabricate at high frequencies. The energy which is not transferred to the other track remains of course on the initial track.

A great variety of ingenious (60), (G4, p.48) uses for multistrip couplers have been proposed but only one or two which could be of importance for the current generation of surface-wave devices will be discussed here.

Bulk waves generated by the input transducer and received at the second transducer form one of the most significant sources of spurious output signals in prototype surface-wave devices, especially when the number of finger pairs per transducer is small. Now if the surface-wave signal is displaced sideways by a multistrip track changer to an adjacent track where it is received by the second transducer, the received signal will be free of bulk-wave clutter because the bulk waves are not coupled by the coupler and remain on the original track (G6, p.172).

The track changer serves another useful purpose in filters using apodized transducers. If a track changer is used between input and output transducer then the beam transferred to the second track is of uniform width at all frequencies despite the

input apodization (G6, p.419), with the result that both transducers can be apodized and the overall response is the simple product of the transfer functions of the two transducers separately. The latter is not true for an in-line apodized pair because the short fingers of the second do not "see" all of the acoustic energy radiated by the first. Thus besides reducing bulk wave interference and reducing diffraction effects from short finger overlaps, the multistrip track changer can offer an extra convenience to filter design.

9.3 Power Divider

Let us now turn to a few examples of the multistrip coupler as a 3 - dB power divider. Here the number of strips per wavelength and the total number of strips are chosen so that the phase shift between modes is $\varphi_L = \pi/2$. Figure 9.3 shows the complex output amplitudes for signals A_1 applied to port 1 on track 1, D_1 to port 2 on track 1, and D_2 to port 4 on track 2. The respective outputs from ports 1 and 3 will be (60)

$$E_1 = (D_1 - j D_2) \sqrt{2}$$
$$E_2 = (D_2 - j D_1) \sqrt{2} \quad . \tag{9.4}$$

Consider three separate origins for the signals D_1 and D_2. If these signals result from two identical reflectors of reflection coefficient r on the two tracks at the same distance from the coupler, and A_1 is the only applied signal, then $D_1 = r A_1/\sqrt{2}$ and $D_2 = -j r A_1/\sqrt{2}$ to within a common phase factor. Sub-

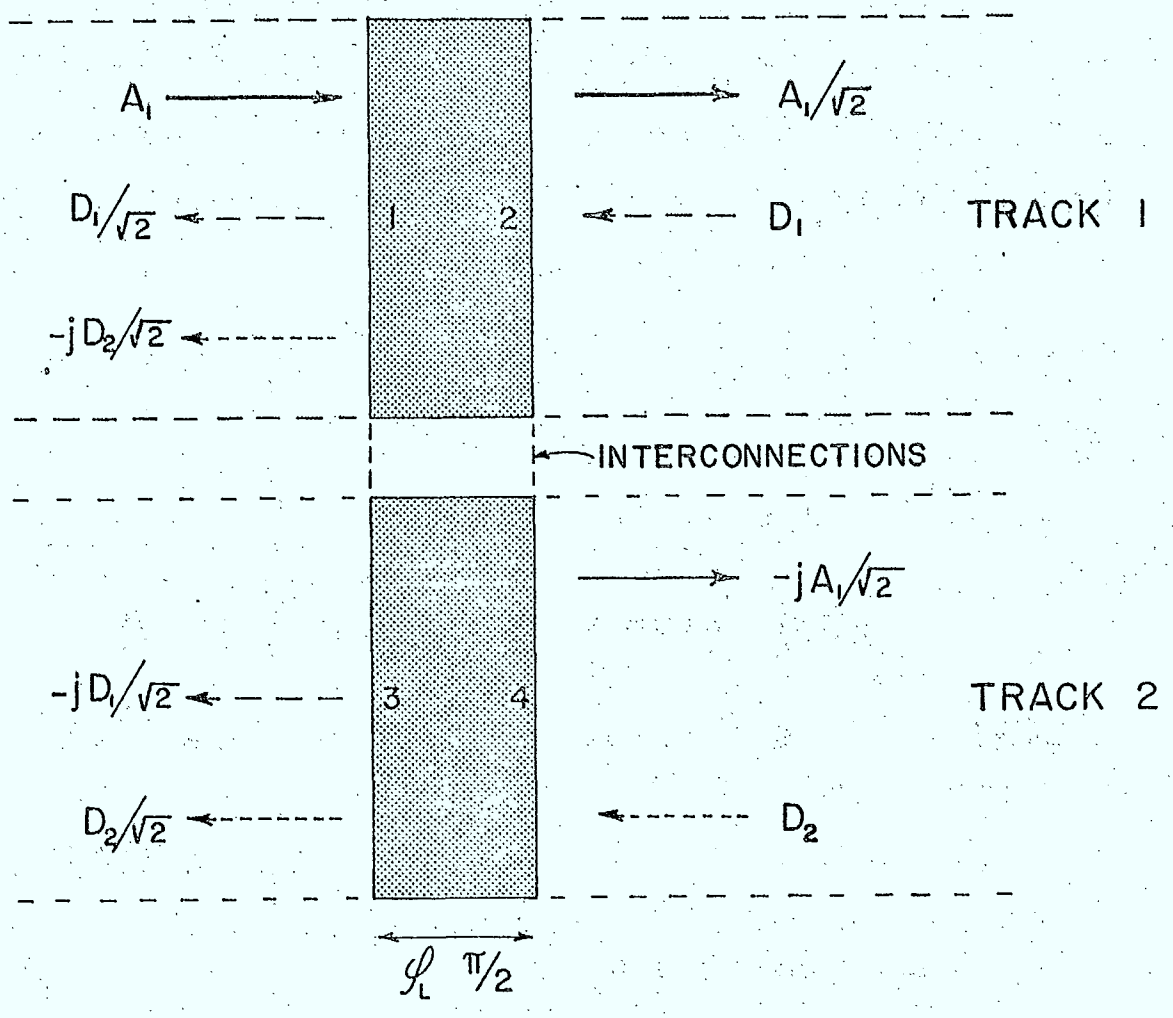


Fig. 9.3 Signal amplitudes in a 3 dB coupler for incident signals A_1, D_1 and D_2.

stituting these values into Eq. 9.4 gives $E_1 = 0$ and $|E_2| = |r A_1|$, thus the only reflected wave appears at port 3 on track 2. This concept can be used to reduce triple transit echoes because if the reflections come from a receiving transducer on track 2 and an identical dummy transducer and load on track 1, the reflections appear only at port 3 and they can be absorbed. The loss in signal would be 3 dB compared to the same geometry with a complete-transfer track changer.

The connections between the strips traversing track 1 and those of track 2 are purely electrical, thus to first order there is no reason to have the strips on track 2 either adjacent or parallel to those on track 1. With this in mind, consider the case of Eq. 9.4 when $D_2 = A_1/\sqrt{2}$ and $D_1 = -j A_1/\sqrt{2}$ which would apply if the lower part of the transducer of Fig. 9.3 were folded back on itself as illustrated in Fig. 9.4. Here, for a given input signal A_1, $E_1 = -j A_1$ and $E_2 = 0$ and thus the component is a broad-band reflector whose reflection coefficient has unity magnitude. Such a component can be used in resonators and with slight modification as a track-changing reflector which allows long delays on a given substrate by reflecting the signal back and forth along successive parallel paths (G10). It should be noted that the strip-by-strip interconnections such as indicated in Fig. 9.4 will not radiate provided the velocity in all other substrate directions is equal to or higher than in the track direction because there is no phase match even if there is piezoelectric coupling. However the capacity represented by the interconnecting part of the strip requires a significant increase in the number of strips for a given fractional energy transfer. Thus such interconnections should be kept as short as the coupler geometry will allow.

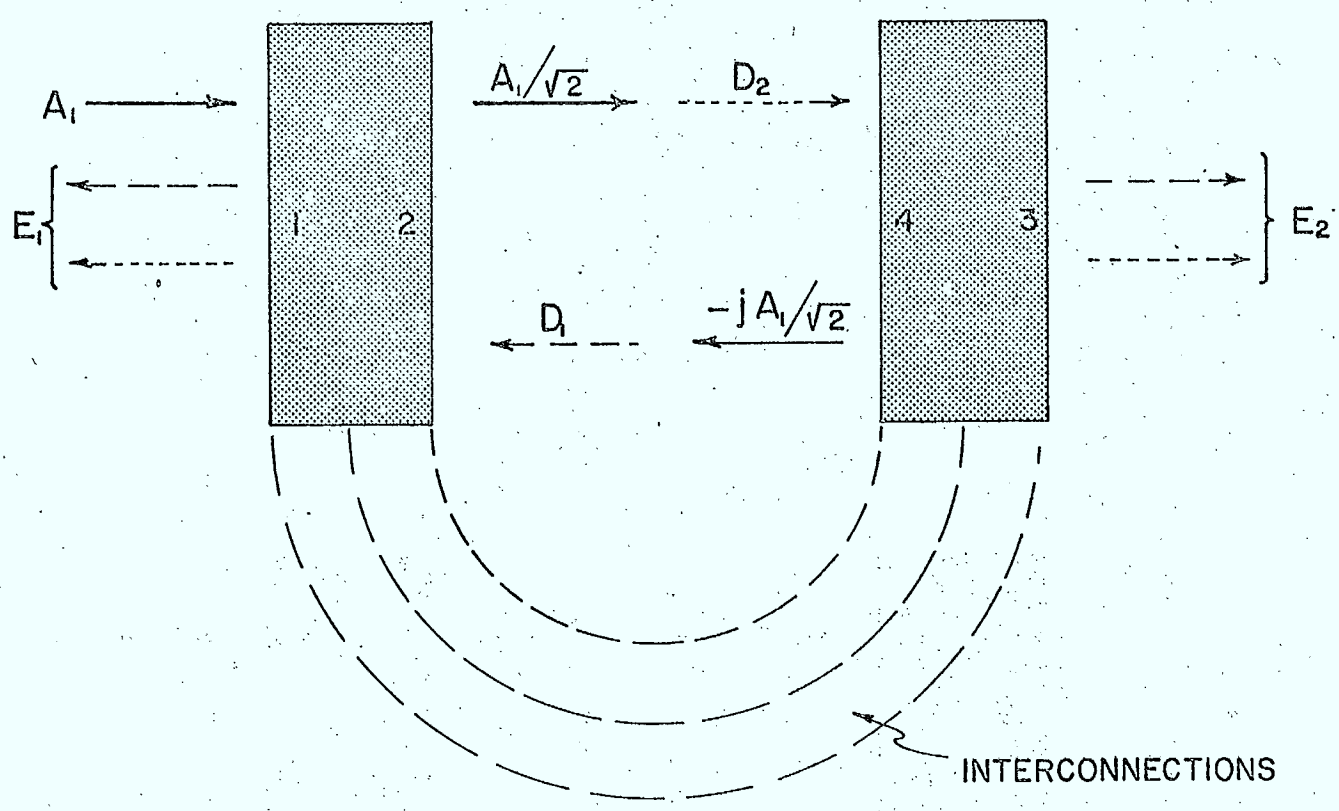


Fig. 9.4 The 3 dB multistrip coupler as a reflector.

As a final example of the multistrip coupler as a power divider (60), consider placing a symmetric uniform transducer offset by $\lambda / 4$ from, or a uniform transducer with an odd number of fingers placed on, the centre line between ports 2 and 4 of Fig. 9.4. Now if A_1 is zero, the input signals will be D_1 and $D_2 = j D_1$ so that from Eq. 9.4, $E_1 = 2 D_1$ and $E_2 = 0$ and thus we have created a unidirectional transducer unit whose bandwidth will be determined by the interdigital transducer itself.

Multistrip couplers can be designed for unequal width tracks, for coupling between different substrates and for performing several functions analogous to those available in electromagnetic waveguide components but their importance in surface-wave devices for systems applications has not yet been definitely established.

X. SUBSTRATE MATERIALS AND FABRICATION

In the course of the discussions of various devices in the earlier sections of this report, it has become evident that the performance of surface-wave components can be critically dependent on specific combinations of the material parameters of the substrate. For example, all the devices contain wavelength-dependent dimensions and thus depend directly on the surface-wave velocity which in turn is determined by the elastic, piezoelectric and dielectric constants of the substrate. Similarly for the structures using interdigital transducers the insertion loss and bandwidth depend directly on the piezoelectric coupling factor which again depends on the material parameters, that is the elastic, piezoelectric and dielectric constants. While the latter fundamental constants are important, the device designer usually is more concerned with the derived parameters such as phase velocity and coupling constant, thus this section will concentrate more on the available values of the derived parameters and their variation from material to material, with angle of propagation on a given surface, and with substrate temperature.

What surface waves usually have to offer to the signal processing system are wide-band devices fabricated with planar technology. Thus for systems applications it is the material performance in the frequency range above some tens of megahertz that is of interest. In this frequency range only single crystals have so far been able to meet the attenuation restrictions imposed by such specifications as dynamic range. The low-loss crystal which has been adopted as a first-choice material against which all others must be judged if piezoelectric coupling is important is $Li\,Nb\,O_3$ and in

particular propagation along the Z direction on a Y - cut face of this crystal. The concentration on this material has been in no small measure a factor in the rapid development of surface-wave device technology and the latter development has in turn resulted in the exponential decrease in price over the past few years of "surface-wave quality" substrates of Li Nb 0_3. This fact is emphasized here because while every free plane solid surface can support a surface-wave, only a very few need be considered in this section on "materials" because only a very few (G4, p.1) offer the combination of properties likely to be required in any surface-wave devices for signal processing.

10.1 Phase Velocity and Electromechanical Coupling Constant

The range of available surface-wave phase velocities on crystalline substrates is from about 1000 to 7000 m/sec. All other factors being equal, lower phase velocities are desirable for applications requiring long delays, but on the other hand higher phase velocities relax the manufacturing tolerances for a given operating frequency. YZ - Li Nb 0_3 with a free surface has a phase velocity of 3488 m/sec while the values for some other common substrates are given in row 3 of Table 10.1 (79, 94).

In surface-wave devices involving piezoelectric coupling with deposited electrodes, the second dominant material parameter is the piezoelectric coupling constant. For example, it was shown in Eq. 3.6 that the input conductance of a uniform transducer at its resonant frequency was equal to

$$G_T(\omega_o) = \frac{8}{\pi} N^2 \omega_o C_L K^2 \qquad (10.1)$$

TABLE 10.1

SUMMARY OF DESIGN DATA FOR MICROWAVE ACOUSTIC SURFACE WAVE DEVICES : BEAM STEERING CALCULATIONS

ASSUME 200 μm TRANSDUCER WIDTHS WHILE DIFFRACTION DATA IS FOR 40 - WAVELENGTH TRANSDUCER AT 1 GHz

1	Material	$LiNbO_3$	$LiNbO_3$	$LiNbO_3$	$LiNbO_3$	$LiNbO_3$	$Bi_{12}GeO_{20}$	$Bi_{12}GeO_{20}$	$Bi_{12}GeO_{20}$	$LiTaO_3$	$LiTaO_3$	$LiTaO_3$
2	Orientation	Y Cut Z Prop	16 1/2° Double Rotated Cut	41 1/2° Rotated Cut X Prop	Z Cut X Prop	X Cut Z Prop	001 Cut 110 Prop	111 Cut 110 Prop	110 Cut 001 Prop	Z Cut Y Prop	Y Cut Z Prop	22° Rotated Cut X Prop
3	Surface Wave Velocity, V (m/sec.)	3488	3503	4000	3798	3483	1681	1708	1624	3329	3230	3302
4	Estimate of Electromagnetic to Acoustic Coupling, $\Delta V/V$	0.0241	0.0268	0.0277	0.0026	0.0252	0.0068	0.0082	0.0037	0.0059	0.0033	0.0027
5	Power Flow Angle ϕ (deg) (Electromechanical)	0	0	0	0	-1.726	0	0	0	0	0	0
6	Temperature Coefficient of Delay $\frac{1}{\tau}\frac{\partial \tau}{\partial T}$ (ppm)	94.	96.	72.	77.	93.	-	-	-	69.	35.	68.
7	S.W. ATTN In Air at 1 GHz (dB/μ sec)	1.07	1.15	1.05	0.93	-	1.64	1.64	-	1.0	1.14	-
8	Air Loading at 1 GHz (dB/μ sec)	0.19	0.21	0.3	0.24	-	0.19	0.19	-	0.23	0.20	-
9	VAC ATTN at 1 GHz (dB/μ sec)	0.88	0.94	0.75	0.69	-	1.45	1.45	-	0.77	0.94	-
10	3 dB Air Prop Loss, Time Delay at 1 GHz, A (μ sec)	2.8	2.6	2.9	3.2	-	1.8	1.8	-	3.0	2.6	-
11	Slope of Electromechanical Power Flow Curve, $\partial\phi/\partial\theta$	-1.083	-1.087	-0.445	+0.192	-0.610	-0.304	+0.366	+0.236	-1.241	-0.211	+0.764
12	Slope of Electromechanical Power Flow Curve, $\partial\phi/\partial\mu$	-0.117	+0.151	0	0	0	0	0	0	+0.556	-0.229	0
13	3 dB Beam Steering Loss, Time Delay, B (μ sec)	7.4	7.0	18.9	46.0	15.8	65.7	53.7	87.6	4.4	15.6	13.3
14	3 dB Diffraction Loss, Time Delay at 1 GHz, C (μ sec)	29.1	29.3	5.1	2.4	7.6	4.1	2.1	2.3	11.3	3.6	1.6

TABLE 10.1 (continued)

SUMMARY OF DESIGN DATA FOR MICROWAVE ACOUSTIC SURFACE WAVE DEVICES : BEAM STEERING CALCULATIONS

ASSUME 200 μm TRANSDUCER WIDTHS WHILE DIFFRACTION DATA IS FOR 40 - WAVELENGTH TRANSDUCER AT 1 GHz

1	Material	Quartz	Quartz	GaAs	GaAs	$Ba_2NaNb_5O_{15}$	Zn O	Zn O	Cd S	Cd S	
2	Orientation	Y Cut X Prop	ST Cut X Prop	110 Cut X Prop	211 Cut 111 Prop	Y Cut Z Prop	Z Cut X Prop	Y Cut Z Prop	Z Cut X Prop	Y Cut Z Prop	
3	Surface Wave Velocity, V (m/sec.)	3159	3158	2822	2621	3177	2716	2665	1729	1716	
4	Estimate of Electromagnetic to Acoustic Coupling, $\Delta V / V$	0.0009	0.00058	0.00008	0.00012	0.0005	0.00707	0.00748	0.00237	0.00262	
5	Power Flow Angle ω (deg) (Electromechanical)	0	0	0	0	0	0	0	0	0	
6	Temperature Coefficient of Delay $\frac{1}{\tau}\frac{\partial \tau}{\partial T}$ (ppm)	-24.	0.	-	-	-	-	-	-	-	
7	SW ATTN In Air at 1 GHz (dB/μ sec)	2.6	3.09	4.22	3.62	~3.7	-	-	-	-	
8	Air Loading at 1 GHz (dB / μ sec)	0.45	0.47	0.40	0.27	-	-	-	-	-	
9	VAC ATTN at 1 GHz (dB /μ sec)	2.15	2.62	3.82	3.35	-	-	-	-	-	
10	3 dB Air Prop Loss, Time Delay at 1 GHz, A (μ sec)	1.2	0.97	0.71	0.83	~0.8	-	-	-	-	
11	Slope of Electromechanical Power Flow Curve, $\partial \omega / \partial \theta$	+0.653	+0.378	-0.537	-2.58	+0.071	-	-	-	-	
12	Slope of Electromechanical Power Flow Curve, $\partial \omega / \partial \mu$	0	0	0	0	0	0	-	0	-	
13	3 dB Beam Steering Loss, Time Delay, B (μ sec)	16.3	28.1	22.2	5.0	149.0	-	-	-	-	
14	3 dB Diffraction Loss, Time Delay at 1 GHz, C (μ sec)	1.7	2.1	6.1	1.50	2.65	-	-	-	-	

and thus proportional to this electromechanical coupling constant K^2. Similarly in Eq. 3.7 it was seen that the maximum bandwidth for a series tuned transducer is obtained when the number of finger pairs is approximately $\pi / 4 K^2$ and the fractional bandwidth is $[4 K^2 / \pi]^{1/2}$. The factor K^2 entered these equations through the equivalent circuit model but it enters other piezoelectric surface-wave coupling phenomena in a similar direct way. For bulk waves K^2 is explicitly defined in terms of the elastic, piezoelectric and dielectric constants of the material and it is found that the fractional change in phase velocity of a bulk wave when the piezoelectric effect is removed (shorted) is $- K^2 / 2$. Thus, by analogy, the factor K^2 is defined as twice the fractional change in the surface-wave velocity v produced by shorting the free surface with a thin conducting film. This quantity

$$K^2 = - 2 (v_{short} - v) / v = - 2 \Delta v / v \qquad (10.2)$$

is calculable from the fundamental material constants and is easily measurable on a given substrate. Other slightly modified definitions are sometimes used, for example (9, 75)

$$K^2 = - \frac{2 (1 + \epsilon_o / \epsilon_s)}{1 - \Delta v / v} \cdot \frac{\Delta v}{v} \qquad (10.3)$$

with ϵ_s given by Eq. 3.2 provides better agreement with experiment of expressions like Eq. 10.1, but in general the simplified expression of Eq. 10.2 is adequate.

For YZ Li Nb O_3, $\Delta v / v$ is .0241 while the values for other materials are given in row 4 of Table 10.1. The tabulated values are calculated (9) from fundamental material constants and are quite sensitive to small errors in the

piezoelectric constants. Measured values of $\Delta v / v$ might be expected to deviate some 10% from these calculated values (75). Li Nb 0_3 provides the highest coupling constant, or $\Delta v / v$, of any of the currently available low-loss crystals.

10.2 Temperature Coefficient of Delay

One of the problems to be faced by the surface-wave component designer is the temperature sensitivity caused by thermal expansion of the critical dimensions and by the temperature dependence of velocity. For most applications the most useful measure of this sensitivity is the first-order temperature coefficient of delay given by

$$\frac{1}{\tau} \frac{\partial \tau}{\partial T} = \frac{1}{l} \frac{\partial l}{\partial T} - \frac{1}{v} \frac{\partial v}{\partial T} \qquad (10.4)$$

where $\tau = l / v$ is the delay time between two points separated by a distance l, and T is the temperature. It will be noted that both thermal effects are incluced in this coefficient because the first term is recognized as the thermal expansion coefficient while the second is the temperature coefficient of velocity.

Values of the temperature coefficient of delay range from -24 to $+96$ ppm $/ {}^{\circ}C$ for the common surface-wave substrates. The importance of this parameter in precision delay lines is obvious, but it is also critical in many other components, for example this coefficient is a useful measure of the degradation of peak-to-sidelobe ratio in pulse compression or spread spectrum systems produced by temperature changes of the surface-wave correlator. In an FM chirp system the change in slope of the time delay versus frequency curve, Fig.5.2a is proportional to this coefficient.

In order to have zero temperature coefficient of delay, the thermal expansion coefficient is required to be equal to the velocity coefficient in the temperature range of interest. This compensation only occurs for rare combinations of crystal, cut and propagation direction. ST cut, X - propagation on quartz has a zero temperature coefficient of delay at room temperature and is thus a very attractive substrate when insertion loss or bandwidth ($\Delta v / v$ is small) can be traded for thermal stability. Other rotated Y - cuts of quartz can be chosen (74, 76) to have zero temperature coefficient of delay above or below room temperature or to have a temperature coefficient to match that of some other part of the system.

Values of the temperature coefficient of delay for a range of crystal cuts are shown in row 6 of Table 10.1. For YZ Li Nb 0_3 this coefficient is relatively high and is one of the few characteristics in which YZ Li Nb 0_3 is inferior to other low-loss piezoelectrics. It has been noted previously that bonds of very close to unity transmission-factor have been made between Li Nb 0_3 and quartz to form a substrate with a single free surface. Such a composite structure with electrodes on the YZ Li Nb 0_3 sections with a temperature coefficient of delay of + 94 ppm / ^{o}C and the central propagation path on YX quartz with its temperature coefficient of - 24 ppm / ^{o}C may be designed to provide a delay line with zero temperature dependence of the overall delay along with the low insertion loss and broad bandwidth of transducers on Li Nb 0_3 .

10.3 Acoustic Attenuation

There is a fundamental loss mechanism for all coherent elastic waves in solids where an inelastic scattering process extracts energy from the coherent wave and gives it to the incoherent lattice waves, thereby the coherent wave is attenuated and the lost energy goes into heat in the solid. Surface-waves also suffer this fundamental form of loss for which the attenuation constant is approximately proportional to the square of the frequency and independent of temperature near room temperature. At high frequencies this propagation loss along with the ohmic losses in the transducer electrodes and any loss due to bidirectionality determine the minimum insertion loss of a surface-wave device. The actual insertion loss depends of course on the matching to the source and load at each frequency of interest.

There is another source of loss which can usually be ignored, this is the attenuation of a surface-wave by coupling to ultrasonic waves in the air in contact with the surface. This air loading has an attenuation constant proportional to the frequency.

The experimentally determined attenuation as a function of frequency is shown in Fig. 10.1 for (a) YX quartz, (b) (110) propagation on a [110] plane of BGO, and (c) YZ Li Nb O_3. These curves lead to the empirical expressions for the attenuation α in dB/μ sec of (79)

$$\text{(a)} \quad \alpha = 2.15\, f^2 + 0.45\, f$$

$$\text{(b)} \quad \alpha = 1.45\, f^{1.9} + 0.19\, f$$

$$\text{(c)} \quad \alpha = 0.88\, f^{1.9} + 0.19\, f$$

where f is in GHz.

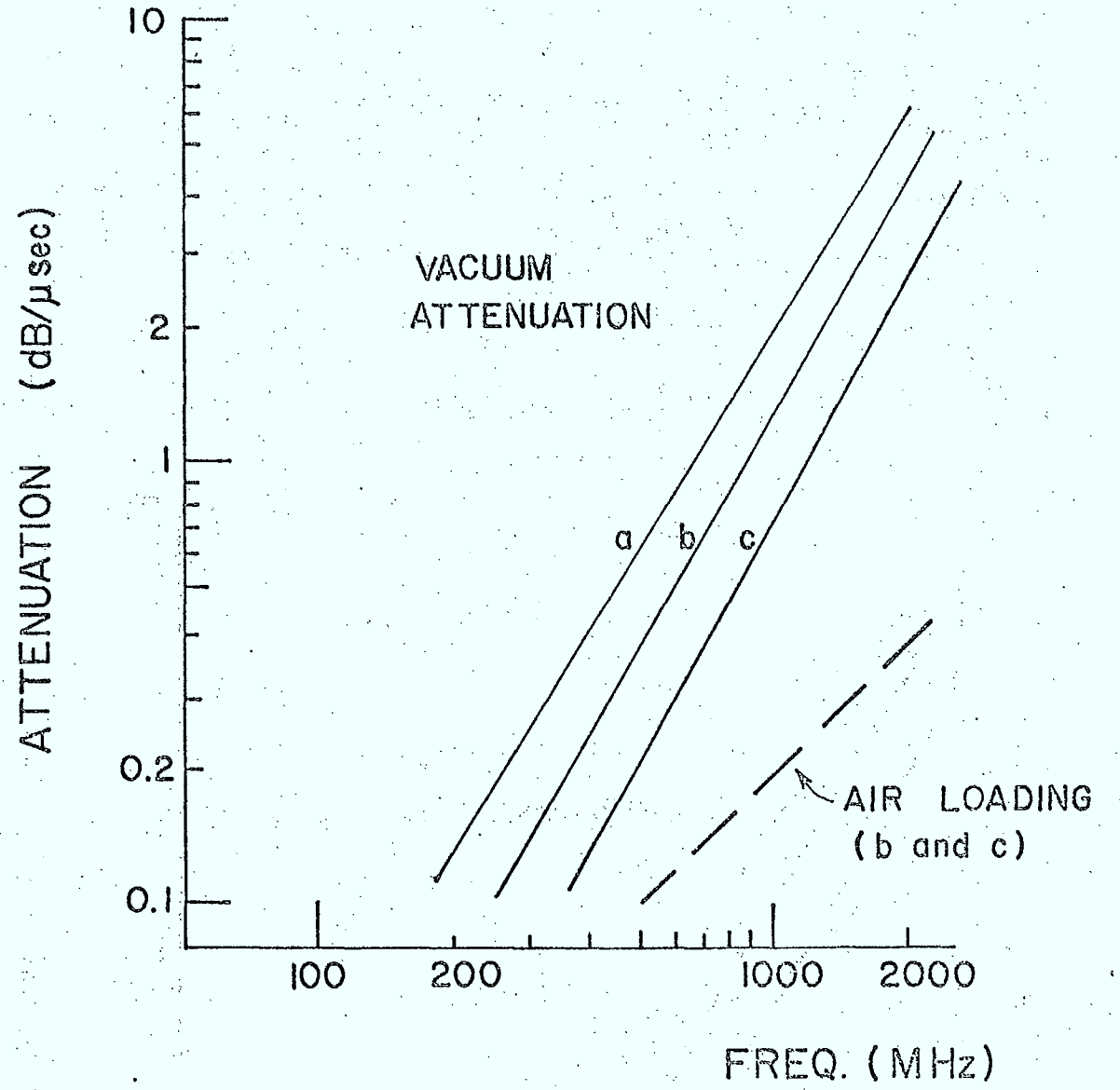


Fig. 10.1 Experimentally determined surface-wave attenuation as a function of frequency for (a) Y X quartz, (b) (110) propagation on a [110] plane of B G0, and (c) Y Z Li Nb 0_3 (79).

It is seen that with these materials, propagation losses are negligible below about 200 MHz unless extremely long propagation paths are involved. Table 10.1 lists in rows 7, 8 and 9 the component and total propagation losses in dB / μ sec at 1 GHz for a range of useful substrates. Scaling to other frequencies as f^2 for the fundamental or vacuum loss and as f for the air loading is usually adequate. Row 10 indicates the propagation loss in another manner, showing the delay time which can be obtained on a given substrate for a loss of 3 dB due to these mechanisms.

10.4 Diffraction and Beam Steering Losses

As discussed in Section 11.1 there is a certain angular spreading of the beam radiated from a transducer due to the finite aperture of the transducer. Most devices operate in the Fresnel region where this diffraction spreading is small, nevertheless because of it some of the acoustic energy radiated from the input transducer is not intercepted by the receiving transducer and the resultant loss of energy due to diffraction contributes to the device insertion loss. Such loss can always be reduced by increasing the aperture size in wavelengths but increases are limited by impedance considerations, apodization ratios, and crystal size. Single crystals are anisotropic and as noted in Section 11.1 the resulting change in velocity with direction can produce either a self-collimating or an increased spreading of the beam radiating from a given aperture. Thus the diffraction loss for a given aperture will depend on the anisotropy of the velocity about the pure-mode direction selected.

On an anisotropic substrate the power flow is colinear with the wave vector only along pure-mode axes. There are several parameters which can be used to measure the anisotropy about a given pure-mode axis, one is $\frac{\partial \varphi}{\partial \theta}$ the change of the angle φ between the power flow direction and the wave-vector direction (row 5 of Table 10.1) with change in the angle θ of the wave vector from the pure mode axis. The slope of the power flow curve $\partial \varphi / \partial \theta = -2b$ in Eq. 11.2. The case $\partial \varphi / \partial \theta = 0$ corresponds to isotropic propagation near the pure-mode axis and $\partial \varphi / \partial \theta \cong -1$ corresponds to large autocollimation (82, 83), that is long time delays with small diffraction loss, Eq. 11.3.

Row 14 of Table 10.1 gives the time delay for 3 dB diffraction loss between two identical transducers 40 λ wide at 1 GHz calculated using Eq. 11.3, ($F = 1.77$) and it is seen that YZ Li Nb 0_3 is extremely self-collimating and the diffraction loss is some ten times less than that of the same aperture in wavelengths on an isotropic substrate. On ST and YX quartz there is more beam spreading than on an isotropic substrate.

Another form of geometrical loss arises in anisotropic crystals because of the limited accuracy with which the geometrical directions can be aligned with the desired crystal axes. Thus while a given choice of propagation conditions, say Z propagation on Y - cut Li Nb 0_3, may represent pure-mode propagation theoretically, there will be errors in the alignment of the crystal face and of the transducers with respect to the desired direction on that face. On an isotropic substrate such alignment is not important but if the crystal is anisotropic near the chosen pure-mode

axis, such errors in alignment will cause beam steering, that is the power will not be colinear with the axis of the transducers and part of the generated beam will not be intercepted by the receiving transducer.

Row 12 of Table 10.1 shows the change in power flow angle per unit change in angle μ of the substrate normal away from the assumed crystal surface, while Row 13 shows the maximum time delay which can be obtained for a 3 dB beam-steering loss due to misalignments $\theta = 0.1^{\circ}$ and $\mu = 0.2^{\circ}$ when the transducer widths are 200 μm each. The beam steering losses are geometrical and thus independent of frequency and proportional to transducer width for a given substrate and alignment errors. The beam steering loss can be reduced by the specification of closer tolerances on crystal alignment.

10.5 Summary of Material Properties

Since YZ Li Nb 0_3 is by far the commonest substrate for surface-wave devices it is useful to compare other possibilities with respect to that ubiquitous cut. Table 10.2 is simply a normalized version of the most important parameters of Table 10.1, the Li Nb 0_3 column is the same whereas all the other columns contain the ratio of each material parameter to the corresponding YZ Li Nb 0_3 parameter and one can see at a glance the relative properties of any of the materials, the last column is for an isotropic material having the same velocity as YZ Li Nb 0_3. For instance, from row 4, the highest coupling material is seen to be X-propagation on the 41.5° rotated Z-cut of Li Nb 0_3. One important feature which is again brought out by this normalization is that although ST quartz has zero temperature

TABLE 10.2

DATA FOR MATERIALS SELECTED FROM TABLE 10.1 RELATIVE TO VALUE FOR Y - CUT, Z - PROP, $LiNbO_3$

GIVEN IN FIRST COLUMN. ROW NUMBERS ARE SAME AS TABLE 10.1

1	Material	$LiNbO_3$	$LiNbO_3$	$LiNbO_3$	$LiNbO_3$	$Bi_{12}GeO_{20}$	$Bi_{12}GeO_{20}$	$LiTaO_3$	$LiTaO_3$	Quartz	Quartz	Isotropic
2	Orientation	Y Cut Z Prop	16 1/2° Double Rotated Cut	41 1/2° Rotated Cut X Prop	X Cut Z Prop	001 Cut 110 Prop	111 Cut 110 Prop	Z Cut Y Prop	Y Cut Z Prop	Y Cut X Prop	ST Cut X Prop	
3	Surface Wave Velocity, V (m/sec.)	3488	1.004	1.147	0.999	0.482	0.490	0.954	0.926	0.906	0.905	1.
4	Estimate of Electromagnetic to Acoustic Coupling, $\Delta V/V$	0.0241	1.112	1.149	1.046	0.282	0.340	0.245	0.137	0.037	0.024	-
6	Temperature Coefficient of Delay $\frac{1}{\tau}\frac{\partial \tau}{\partial \tau}$ (ppm)	94	1.021	0.75	0.989	-	-	0.734	0.372	-0.255	0	-
10	3 dB Air Prop Loss, Time Delay at 1 GHz, A (μ sec)	2.8	0.929	1.036	-	0.643	0.643	1.071	0.929	0.429	0.346	-
13	3 dB Beam Steering Loss, Time Delay, B (μ sec)	7.4	0.946	2.554	2.135	8.878	7.257	0.595	2.108	2.203	3.797	∞
14	3 dB Diffraction Loss, Time Delay at 1 GHz, C (μ sec)	29.1	1.007	0.175	0.261	0.141	0.072	0.388	0.124	0.058	0.072	.097

coefficient of delay, its diffraction-loss time-delay is inferior to an anisotropic material as can be seen from row 14 . BGO with phase velocities half those of Li Nb 0_3 is attractive for long delay applications, and in addition the lower coupling constant and increased diffraction loss may be offset by the lax tolerances on crystal alignment (row 13) .

Layered Structures

It is often necessary to deposit piezoelectric films onto certain non-piezoelectric substrates in order to exploit some specific desirable property of the substrate. For example the use of sapphire as a substrate which is compatible with silicon-on-sapphire technology thus making it possible to marry acoustic-surface-wave and silicon-on-sapphire technologies.

Zinc oxide has been successfully sputtered on oxidized silicon and on fused quartz (28) and the best results reported are shown in Fig. 10.2 which gives calculated and measured values of coupling factor for these films as a function of film thickness for thicknesses between 2 and 4.5 μ . The calculated K^2 curve is for C - axis oriented single crystal Zn 0 on fused quartz. The highest measured K^2 value of 2.4 % for 0.35 wavelength thickness films is only 15 % less than theoretical.

Aluminum nitride films have been deposited on sapphire by chemical vapour deposition. Current work (17), (G5, p.274) has used Al N films about 1.6μ thick, thicker films tend to have an irregular surface which require mechanical polishing before any transducers can be formed on the surface. The best values of

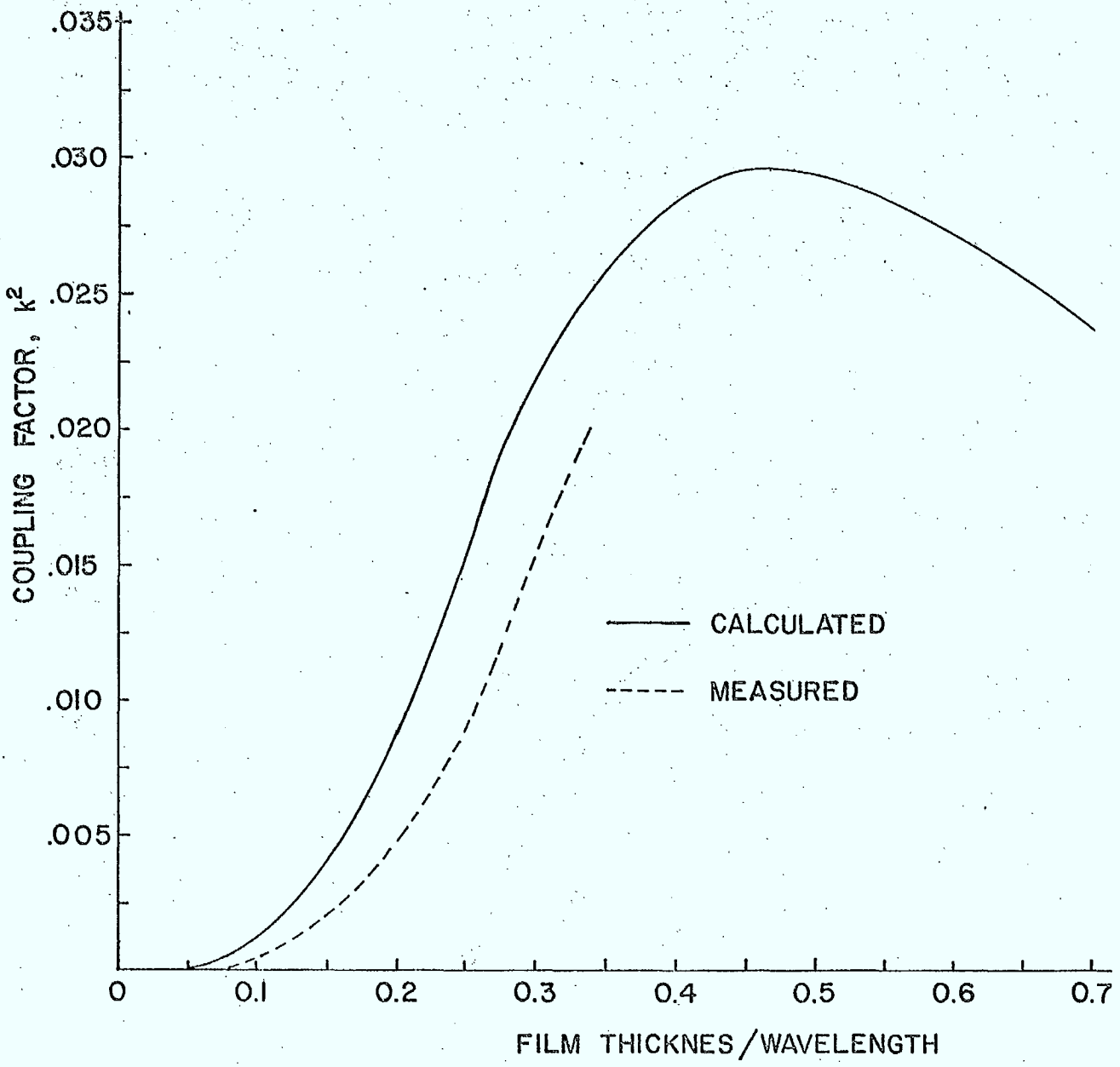


Fig. 10.2 Calculated and measured coupling factor, K^2, for Zn 0 layers on fused quatrz and oxidized silicon as a function of thickness / wavelength ratio (28).

K^2 obtained are 0.6 % for 0.1 λ thick films. Thicker films will be required to get maximum coupling and hence widest bandwidth at minimum insertion loss. These thin films have been successfully used to construct surface-wave programmable matched filters (G6, p.333) in a side by side configuration with silicon-on-sapphire switching circuits on a common sapphire substrate.

Non-piezoelectric Materials

It is obviously necessary to know the acoustic properties of materials used as substrates in layered structures or materials used in the bonded composite structures described earlier. In Table 10.3 a small selection of non-piezoelectric materials has been assembled and the surface-wave velocity for specific directions given. Although Te 0_2 is, strictly speaking, a piezoelectric, its coupling constant is so small that it can be considered negligible ; it has been included in Table 10.3 because it has a zero temperature coefficient of delay and has such a low velocity that it is potentially useful in composite structures.

10.6 Fabrication Techniques

In this and the remaining sections of this chapter, the techniques used in the fabrication of surface-wave devices such as the interdigital transducer are reviewed and compared with the planar fabrication technology used for integrated circuits. The specific aspects of the technology discussed include sample preparation, thin film deposition and metallization, pattern definition techniques including artwork and mask generation, photolithography, electron beam and x-ray processes and etching and

TABLE 10.3

SUMMARY OF DATA FOR SURFACE WAVES ON SELECTED NON-PIEZOELECTRIC MATERIALS

1	Material	Rutile	Rutile	Sapphire	Silicon	$T_e O_2$	YAG	YAG	YIG	YIG
2	Density (Kg/m^3)	4260	4260	3986	2332	5990	4550	4550	5170	5170
3	Orientation*	Y Cut Z Prop	Z Cut X Prop	Z Cut Y Prop	Z Cut X Prop	58.2° Cut X Prop	Z Cut X Prop	$\bar{1}10$ Cut 001 Prop	Z Cut X Prop	$\bar{1}10$ Cut 001 Prop
4	Surface wave Velocity (m / sec)	4194	4789	5706	4921	1387	4599	4603	3584	3579

* Orientation is given in terms of Crystal axes, Miller indices or Euler angles.

engraving methods. Since interdigital transducers play a key role in devices special attention is given to their fabrication.

Substrate Preparation

As discussed in the previous sections devoted to materials, surface-wave substrates are usually single crystals of Li Nb 0_3, quartz etc. In contrast to micro-electronics, surface-wave substrates require precise orientation with respect to crystalline axes in order to operate along pure-mode axes and minimize undesirable beam steering effects, thus skilled and lengthy procedures are involved in cutting and polishing the surfaces to very stringent optical quality specifications. Typically, as discussed in Section 10.4, the crystal cut and orientation has to be within 0.1° of specification to minimize diffraction and beam steering losses.

The substrates are mechanically polished to give a finished surface quality of one tenth of optical wavelengths with negligible surface damage to reduce scattering losses. Note that the high degree of precision required with respect to crystalline orientation coupled with non-standard sizes results in a very high cost for the substrates (G4, p.22) when compared to a high degree of standardization in the semiconductor industry.

10.7 Film Deposition and Metallization

Transducer fabrication requires the deposition of thin metallic films on the substrate and similar metal films are also used in waveguiding applications and in multistrip coupled devices. Thin films are also used to achieve prescribed dispersive

characteristics, control temperature coefficients of delay, obtain electroacoustic interactions with carriers, and to excite surface-waves on non-piezoelectrics using piezoelectric films.

Prior to any deposition, the substrates are prepared by removing any undesirable surface impurities by etching and cleaning in organic solvents (acetone, methanol, TCE) and if needed the substrate is coated with photoresist for either a lift-off or stripping process to be described in the next section.

Thin films are deposited by thermal evaporation or electron beam methods, or by DC / RF sputtering. Which method is used is dependent more often than not on the availability of a suitable vacuum system. Gold and aluminum are the most common metals used for films ranging from a few hundred to a few thousand angstroms. Sputtering is in many respects preferable to evaporation since better adherence and uniformity is achieved over large areas, properties that are important for good edge definition in the lithographic process. With gold it is usually necessary to deposit an underlayer of chromium or titanium to get good adhesion on oxide substrates. As discussed in the sections on materials, zinc oxide is used as a piezoelectric thin film for transduction on non-piezoelectric substrates and has been successfully deposited (28) by DC and RF sputtering in a reactive atmosphere (Argon + Oxygen) yielding good quality reproducible films. Aluminum nitride has been deposited on sapphire by thermal evaporation (102), RF sputtering (78), and by chemical vapour desposition (68). The latter is a very promising process which can be coupled with silicon-on-sapphire technology on a common substrate for the fabrication of programmable devices (Section 6.3).

10.8 Pattern Definition

The main difference between surface-wave devices and microelectronics is in the pattern generation and replication techniques. Patterns for surface-wave devices are usually much larger in size ($\simeq 1"$) but require higher resolution ($< 1\ \mu m$) than their microelectronics counterparts, are usually only single level but require alignment with crystalline axes to within 0.1°.

Photoresist Contact Printing

The process of contact printing is shown schematically in Figs. 10.3 and 10.4 illustrating the photomask production, chemical etching, and lift-off or stripping techniques.

The photomask is usually produced in two stages, first the high resolution parts of a pattern such as the transducer fingers are produced on a photographic emulsion reduced by factors of the order of 10 x by direct photography of the artwork on the Rubylith, these sub-patterns are then assembled on a second sheet of Rubylith on which coarse resolution parts such as contact pads have been cut and a final reduction is carried out producing the final photomask on a high resolution plate. Alternately a step and repeat camera is used to produce the required pattern, this provides better alignment and precision.

The photolithography is done in one of two ways (3, 84) either the chemical etching process or the resist stripping (lift-off) process. In the chemical etching process the metal film deposited on the substrate is first coated with photo-resist such as Kodak - KTFR (negative) or Shipley AZ 1350 (positive), negative or

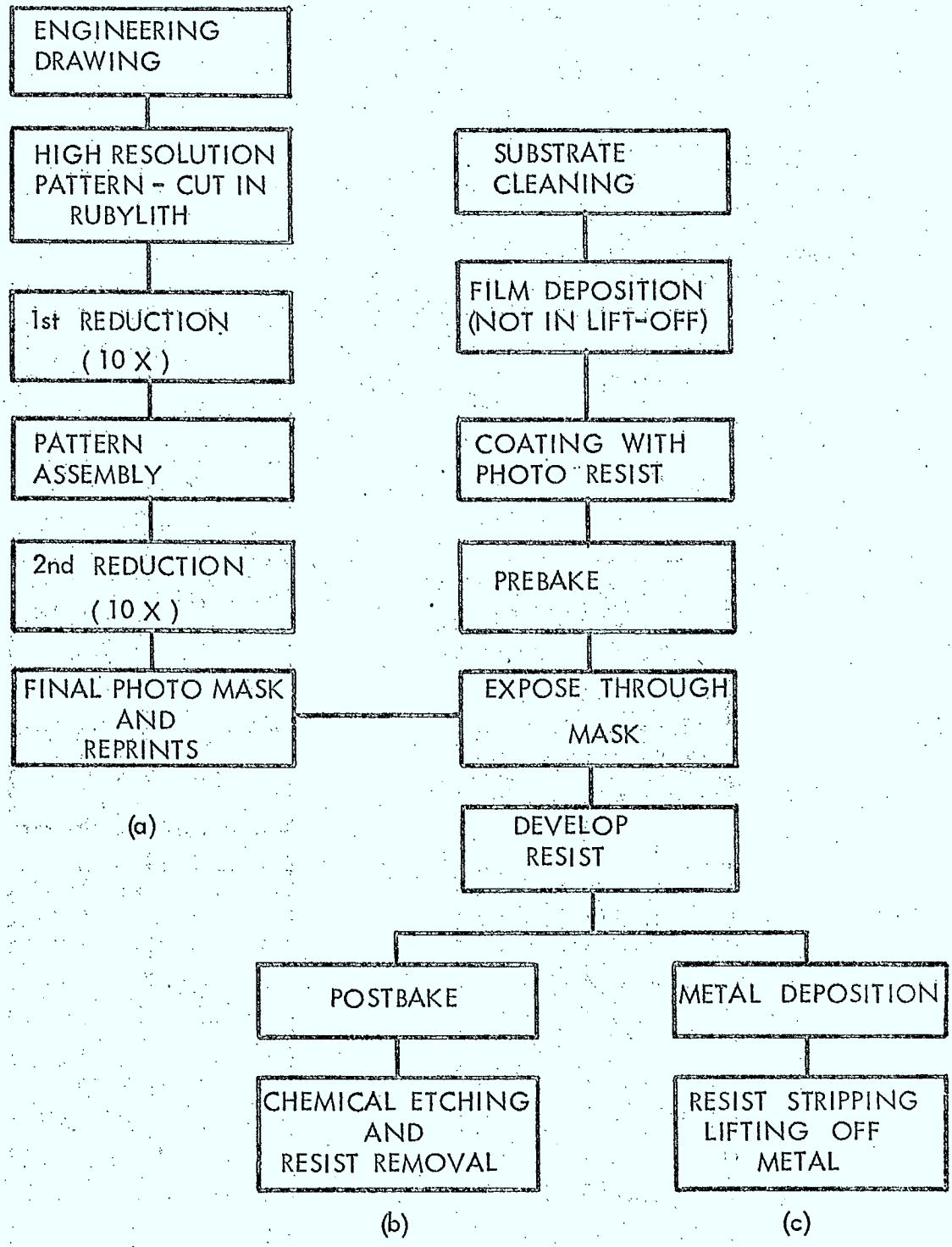


Fig. 10.3 Photo-lithographic contact print processes.

(a) Artwork and photomask production, (b) Chemical etching process, and (c) Lift-off or stripping process (G4, p.22).

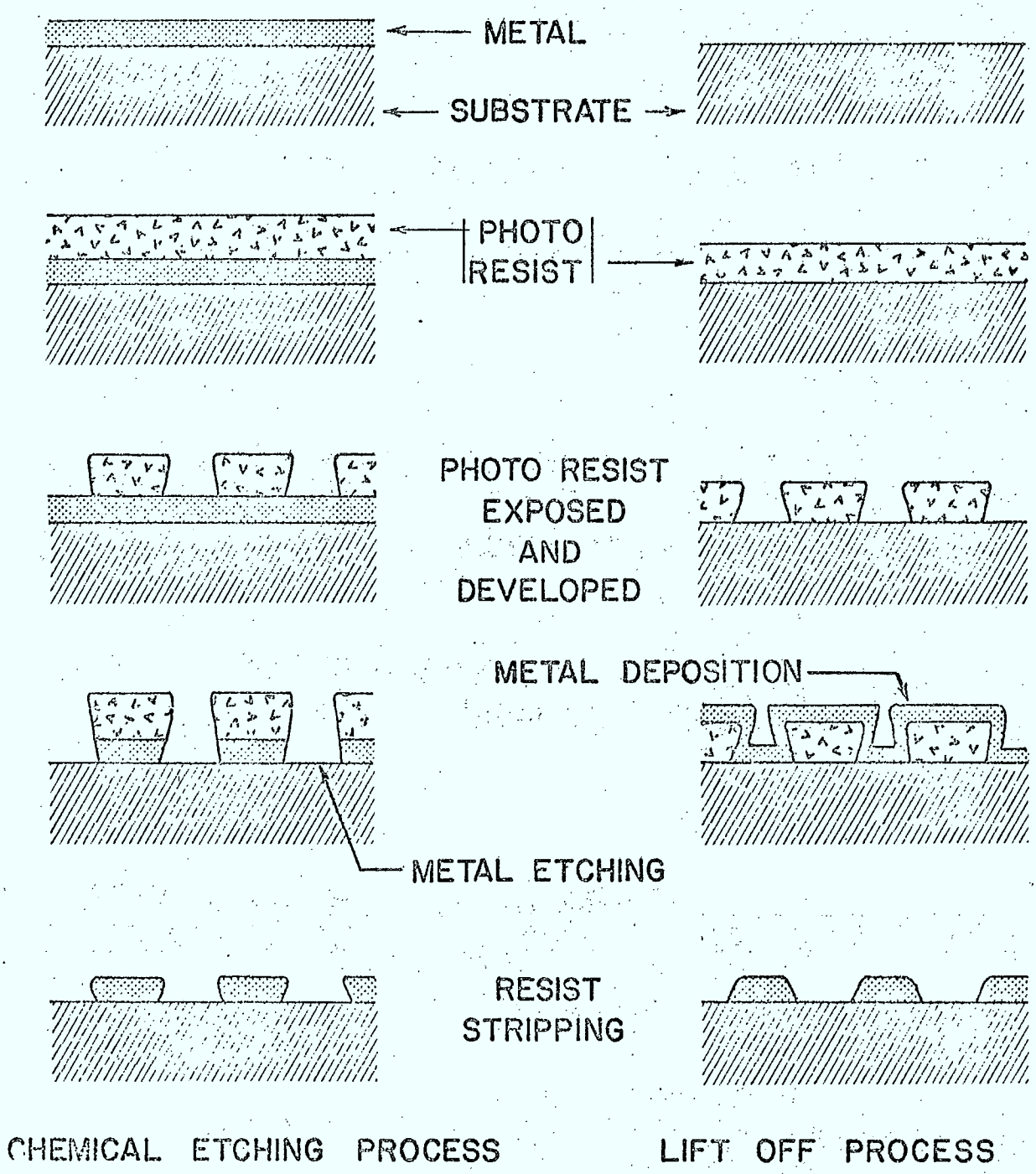


Fig. 10.4 Schematic representation of photo-lithographic processes.

positive meaning that the developing solvent removes the unexposed photoresist or the exposed polymer respectively. Spin coating gives the best uniformity of photoresist coating for small samples. For rectangular substrates having large aspect ratios spinning causes uneven coating and dipping with slow withdrawal (44) is more suitable.

After resist coating, a prebaking period at a slightly elevated temperature (20 minutes at 70° C for AZ 1350) is needed to cure the photoresist. The pattern is then produced in the photoresist by contact printing using ultraviolet rich light, exposure times depend on resist thickness, light intensity and pattern resolution and can be from 15 to 150 secs. Degradation of resolution is avoided if the mask is kept in intimate contact with the substrate and to this end "conformable masks" made of an 800 A thick chromium pattern on Corning type 0211 glass, 0.2 mm thick (84) have been used to improve replication accuracy and definition. This is possible because the thin glass is sufficiently flexible to conform to the substrate contours when used in a vacuum chuck (106) such as is shown in Fig. 10.5 .

After exposure the pattern is developed using a developer appropriate to the photoresist used. The time, temperature and agitation of the developer bath depend on photoresist thickness (44) and a postbaking of the developed resist is often required (20 minutes at 90° C for AZ 1350) for high resolution patterns.

Chemical etching is now carried out as illustrated in Fig. 10.4. Aluminum is etched in 20% Na OH or KOK or in a H C l bath. Chromium-gold requires two etchants, acqua-regia or iodine-potassium iodide for the gold, ferricyanide-

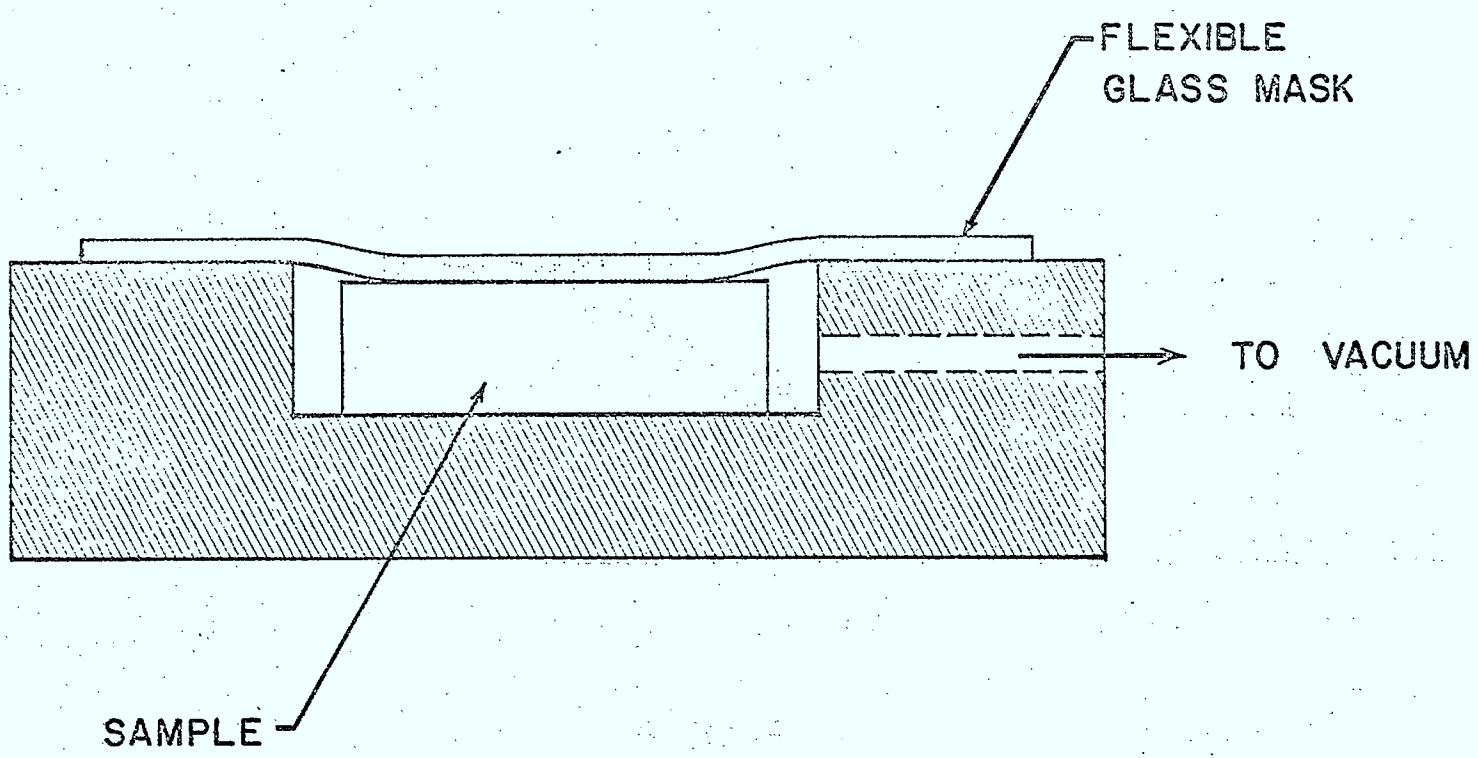


Fig. 10.5 Diagram of vacuum frame for use with flexible photomasks.

hydroxide mixtures (3) for the chromium. The choice of etchant is very difficult since it is necessary to avoid chemical damage of the substrate and also prevent damage to the photoresist when etching the metal.

The lift-off or stripping processes as depicted in Figs. 10.3 and 10.4 is similar to the chemical etching method except for the film deposition which is done onto the developed photoresist pattern. This pattern is stripped (with acetone spray for AZ 1350) thus lifting off the deposited metal from the regions initially covered with resist; the lifting of metal is initiated at the substrate-metal-resist areas. The difference in the final result is mainly the elimination of undercutting which occurs with chemical etching (see Fig. 10.4) and limits the resolution of the latter method to line widths of $\sim 5\ \mu m$ whereas the lift-off process can achieve line resolution of at least $1\ \mu m$.

Noncontact Printing

The resolution limits in the preceding methods are primarily determined by light scattering in the resist material during exposure. Electron beam lithography has been used (8, 73) to overcome these limitations and to extend the line resolution capabilities and hence the operating frequencies of surface-wave devices. Electron beam lithography utilizes an electron-sensitive resist illuminated with electrons to produce the pattern. Patterns can be produced by electronically scanning the surface of a resist-coated substrate, thereby writing the pattern on the surface. A computer controlled Scanning Electron Microscope (SEM) is used to generate and register the pattern. This system is illustrated in Fig. 10.6, the electron beam diameter is

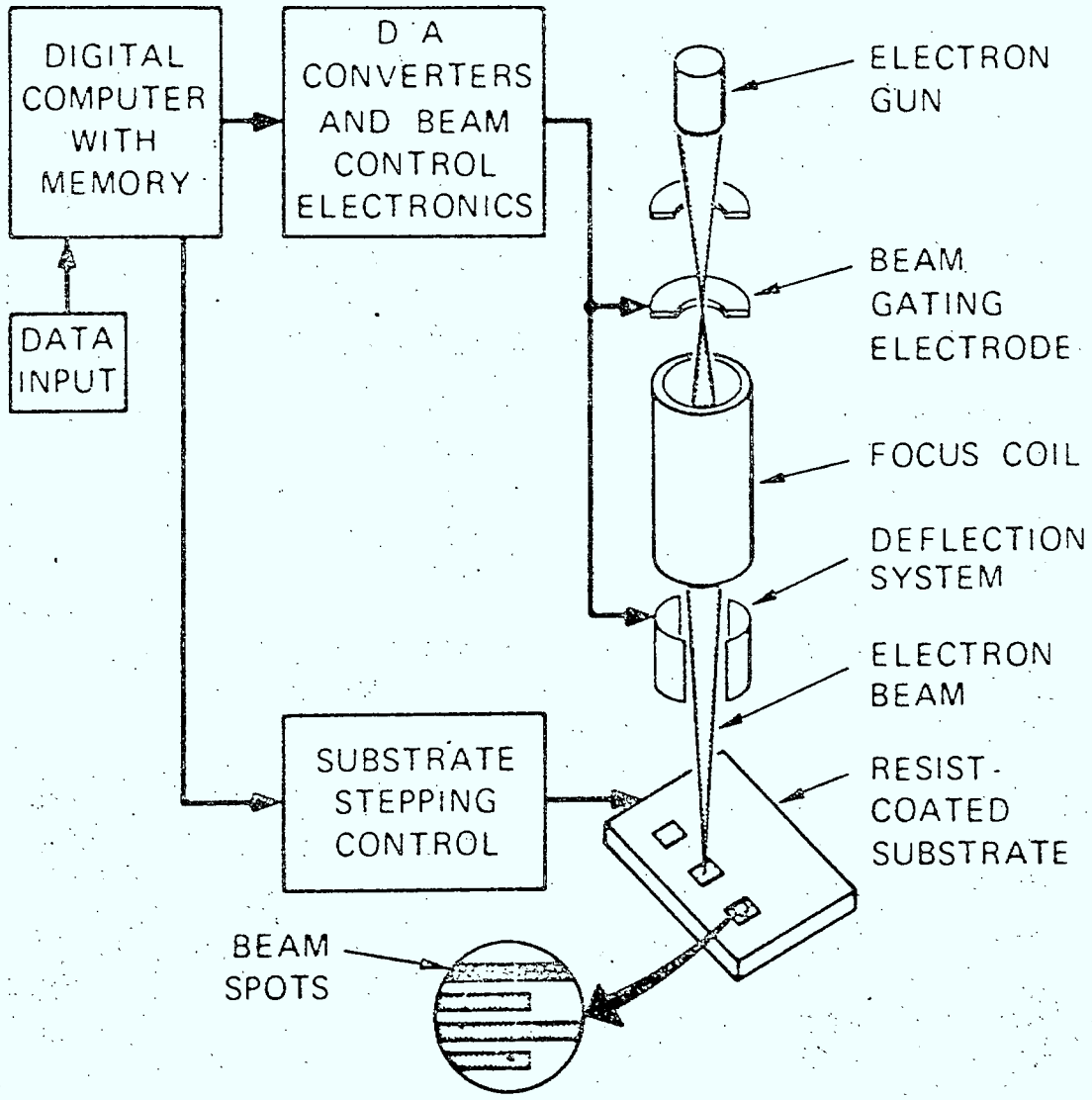


Fig. 10.6 Electron beam lithography using a scanning electron microscope (G4, p.22).

typically 100 A°, thus indicating the degree of resolution ideally possible. Electron scattering in the resist limits the practical resolution of electron beam lithography to 0.1 μ m (G4, p.22). Since the field of view in a SEM is typically 100 μ m, very accurate step and repeat mechanisms using laser interferrometric techniques are required for patterns in excess of 100 μ m .

An alternate method is to project onto the surface an electron image through a mask or series of masks. Such a scheme is illustrated in Fig. 10.7 which was designed specifically to produce interdigital patterns for surface-wave devices. In this scheme, a slit lens images a hole pattern into a series of lines and has the advantage that some demagnification is used, namely the hole sizes in the object pattern are 100 times the line widths in the final image. Thus the hole pattern can be produced by conventional photolithography and at present 0.3 μ m lines over an area 0.5 x 0.5 m can be produced.

In view of the expense and low throughput associated with the scanning electron lithography technique, this method is being used primarily for mask making, and device fabrication is carried out using a high-resolution contact-printing or projection technique.

Another technique capable of similar resolution is the x-ray lithographic (89) process shown in Fig. 10.8. Here a resist coated surface is exposed through a mask to soft x-rays ($\lambda \cong 10$ A). Since diffraction of such wavelength radiation is negligible the mask and substrate may be separated by an amount limited only by the size of the source and the target to substrate distance.

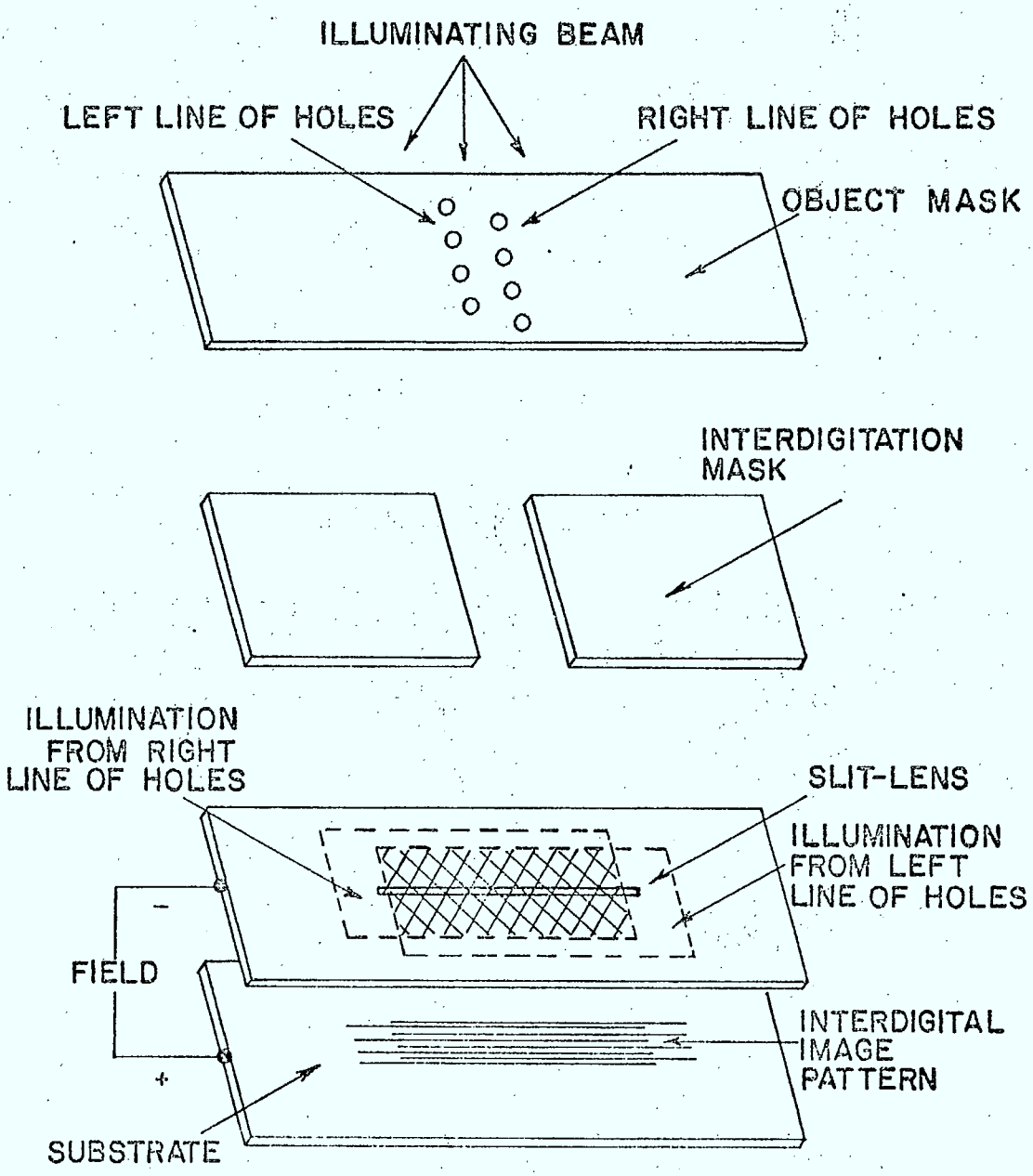


Fig. 10.7 Diagram of electron beam projection lithography system (G4, p.22).

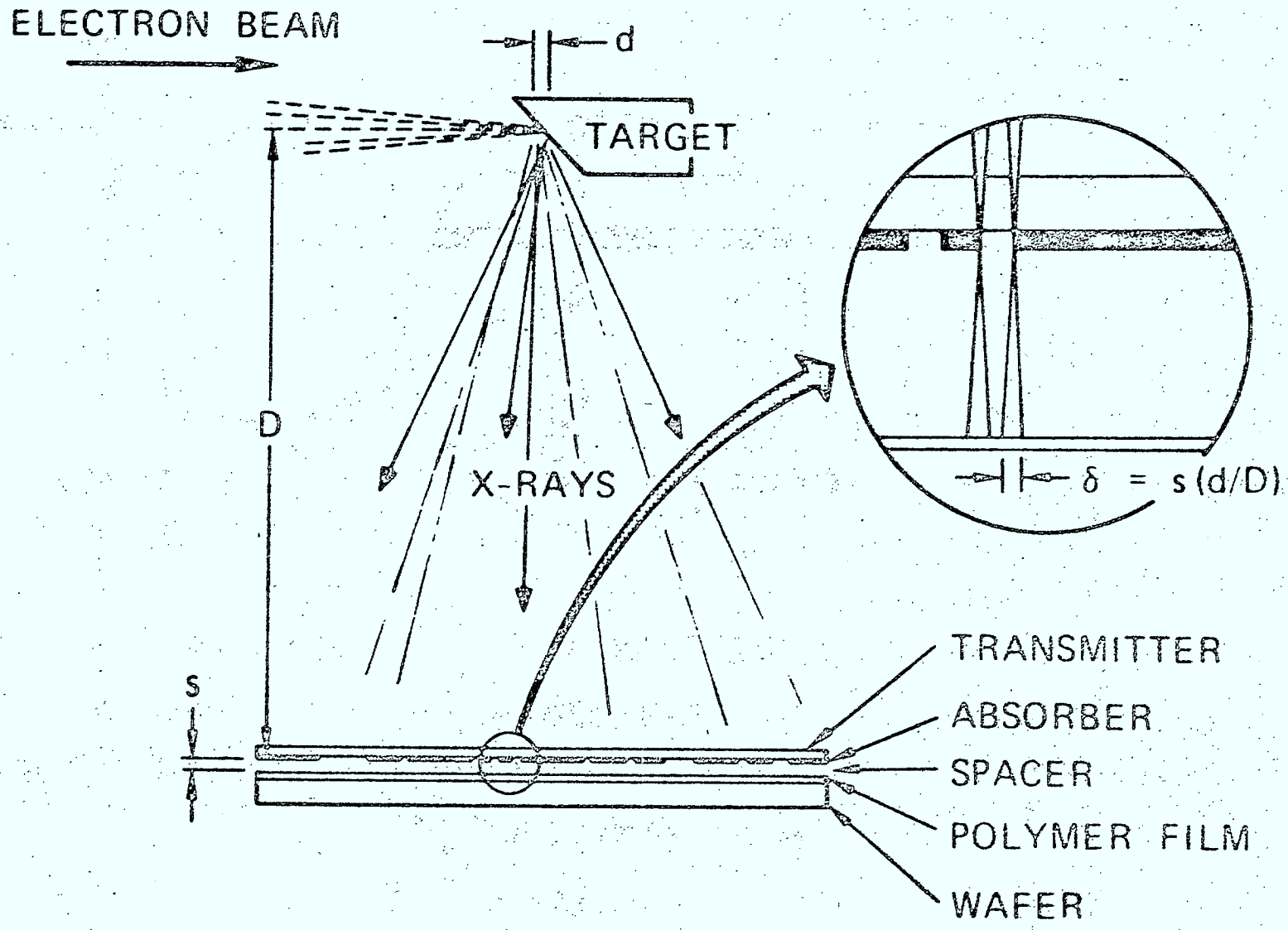


Fig. 10.8 Diagram of soft x-ray lithographic system. Inset shows image blurring due to finite size of x-ray source (G4, p.22) .

10.9 Etching and Engraving Processes

In the fabrication of devices such as reflective array compressors a pattern is etched or engraved in the substrate material (106). The fabrication procedure consists in registering the pattern in photoresist on the substrate by a standard photolithographic technique which eventually leaves a resist mask covering the regions which are not to be etched. The etching can be achieved either by ion-beam bombardment or sputter etching of the substrate. The protective resist coating must be thick enough to prevent the removal of all the material in the covered regions.

One of the reasons which explain why surface-wave devices have reached the practical application phase so quickly is that they use the same planar techniques used for microelectronic devices. A number of differences become apparent when one considers the large scale production of surface-wave devices. A comparison of the fabrication aspects for these two types of devices is shown in Table 10.4 which is taken from an excellent review paper by Bahr (G4, p.22).

TABLE 10.4

COMPARISON OF FABRICATIONAL ASPECTS

FOR SURFACE-ACOUSTIC-WAVE DEVICES AND MICROELECTRONICS

	Surface-Acoustic-Wave Devices	Microelectronics
Substrates	Produced in small quantities ; expensive ; non-standard sizes.	Inexpensive and mass-produced ; standard size.
Polishing	Mechanically polished ; surface quality may vary by a factor of 10 from batch to batch ; invisible surface damage.	Chemically polished ; smooth and uniform.
Crystalline Orientation Relative to Reference Edge on Substrate	Alignment to within 0.1 often required in plane of propagation.	Alignment to within 1 in plane of wafer often adequate.
Mask Alignment	Usually single level ; pattern on mask often required to be aligned to a crystalline axis to be within 0.1	Usually multi-level ; alignment of patterns and crystalline axes usually not critical.
Device Size	Patterns are repetitive and cover large areas ; the entire pattern usually must be defect-free ; "chip" size may be several inches in extreme cases.	Patterns are small ; "chip" size is of the order of 0.1 or 0.2 inches many devices are fabricated on a single wafer.
Photomask	Large-area patterns often require photo composition at the reticle (10x) or final size level ; very accurate control of step-and-repeat is required.	Entire pattern can usually be photo-reduced from an oversize drawing.
Metalization	500 - 5000 Å thick ; usually aluminum.	Usually 5000 Å or thicker ; both gold and aluminum used.
Etching	Chemical etching adequate for low-frequency devices (10 - 200 MHz) : "lift-off" techniques more promising for higher frequencies and large - area devices.	Mostly chemical etching.

XI. DIFFRACTION, GUIDING AND LONG DELAYS

11.1 Diffraction

It has been mentioned several times previously that the spreading of the acoustic beam between transducers due to diffraction is of importance in surface-wave devices. Here we wish to consider such diffraction effects somewhat more quantitatively.

The situation is analogous to the classical optics problem of radiation from a slit. In the surface-wave case the aperture is effectively the width of the transducer, typically ten to one hundred wavelengths wide, and scalar diffraction theory can be applied taking say one component of the surface displacement as the scalar quantity involved. In the case of long transducers there is some ambiguity in the location of the effective aperture. While patterns can be calculated by adding the contributions to the field pattern from each individual finger pair it is usually adequate to assume that the aperture location is the centre line for symmetric transducers and at the location of the effective fingers for dispersive transducers (38).

Thus at any one frequency there is two-dimensional scalar diffraction with a uniformly illuminated aperture of effective width $2w$ radiating into a half-space, the free surface in front of the transducer. If the substrate is isotropic, the diffraction patterns are those of the slit in scalar optics and when a dimensionless parameter

$$F = \frac{\lambda}{2w} x \qquad (11.1)$$

is introduced as a measure of the distance x from the aperture, there is a cross-

sectional profile characteristic of each value of F as illustrated in Fig. 11.1. For values of $F < 1$ that is in the so-called Fresnel region the radiation pattern is a more-or-less parallel beam with ripples across the amplitude profile but with most of the energy contained in a strip equal in width to that of the aperture and the phase varying little ($< \pi / 2$) over the cross-section. Values of $F > 1$ correspond to the Fraunhofer region in which the angular pattern remains constant as the distance from the aperture is increased and the phase becomes that of a circular wavefront centred on the aperture.

Most surface-wave devices operate in the Fresnel region because it is desired to have the output transducer intercept most of the radiated beam and, most important, to have the phase of the intercepted beam constant along the length of the transducer fingers. Note that at $f = 100$ MHz and $\lambda = 30\ \mu$ for an aperture $2w = 2$ mm $= 66\lambda$, $F = 0.5$ corresponds to a distance of 66 mm from the aperture.

While the axis between transducers is almost invariably a pure-mode axis when the surface-wave device is fabricated on a single-crystal substrate, directions away from this axis are not in general pure-mode directions because of the crystal anisotropy. For such directions, energy flow is not parallel to the wave vector. The method of calculation of the radiation patterns for this case (38, 61, 62, 79, 93) will not be discussed here but one or two important quantitative features can be noted. The anisotropy is usually sufficiently small that the surface-wave velocity as a function of the angle θ measured from the pure-mode axial direction where the velocity is

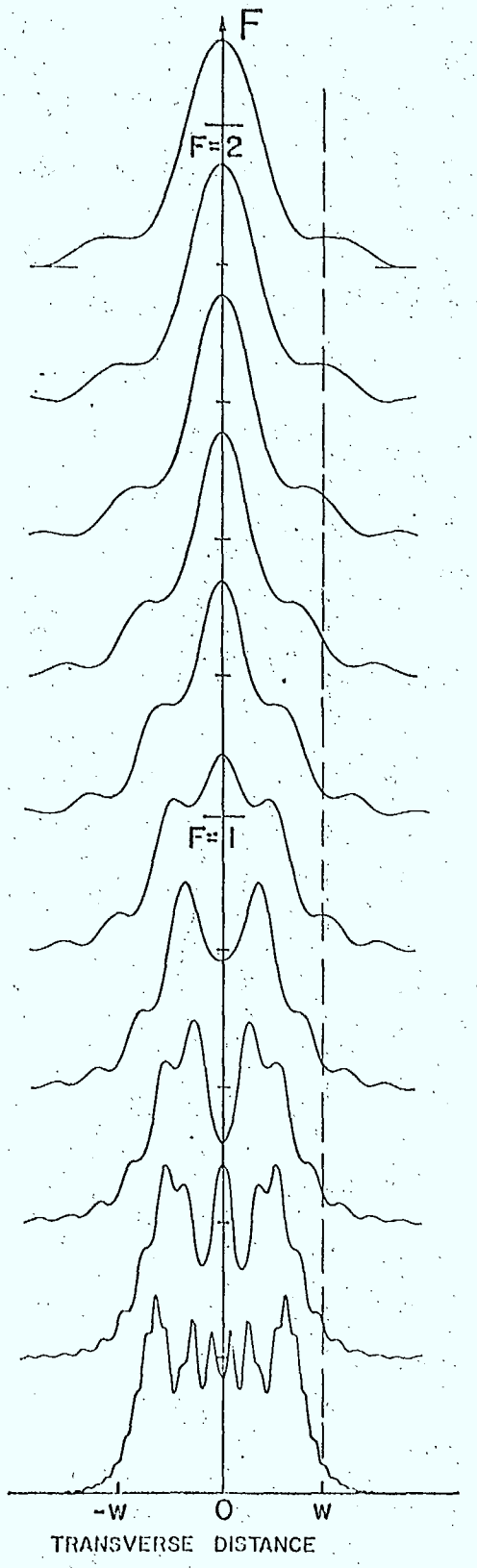


Fig. 11.1 Transverse diffraction profiles for different values of F, Eqs. 11.1 and 11.3 (93).

v_o can be represented by

$$v = v_o (1 - b \theta^2) \qquad (11.2)$$

the factor b is a material parameter dependent on the choice of crystal, cut and axis; typical values are given in Table 10.1 where

$$b \cong - \frac{1}{2} \partial \varphi / \partial \theta \ .$$

It can be shown that in the parabolic approximation of Eq. 11.2, the diffraction profiles at successive values of F are the same as those of Fig. 11.1 provided F is modified from Eq. 11.1 to (38)

$$F = \frac{\lambda}{w^2} (1 - 2 b) x \quad . \qquad (11.3)$$

Positive values of b result in a given profile occurring further from the aperture than in the isotropic case, and in particular the length of the Fresnel zone becomes longer. As b approaches 0.5, the length of the Fresnel zone approaches infinity, that is the aperture distribution of intensity is maintained out to large values of x and we approach an autocollimated beam. In Section 10.4 it is noted that YZ Li Nb 0_3 approaches the autocollimating condition, though for this geometry the parabolic approximation is not very valid. The ST geometry on quartz has a negative value of b and there is more beam spreading than on a similar isotropic substrate. Table 10.1 gives some estimates of the diffraction loss for typical cases and also the time delays which can be achieved within a specified diffraction loss. Note that for a pair of transducers of the same width 2 dB diffraction loss occurs when F = 1.77 (79) .

11.2 Waveguiding

Surface-waves are guided in one dimension by the free surface itself and it is possible by several means to restrain the wave in the other dimension perpendicular to the direction of propagation. Thus one can produce the analog of the guided wave in the electromagnetic case and conceptually at least, and in most cases experimentally, one can design the analogs of all the myriad components of electromagnetic waveguide plumbing (92). While a lot of theoretical and preliminary development work has been done on waveguides for acoustic surface-waves (G10, 45, 70, 98, 100), such components have not been used commercially as yet, in large part because no system need really requiring them has appeared, and in part because they have tended to be lossy, dispersive or difficult to fabricate.

Three different forms of waveguide are illustrated in cross-section in Fig. 11.2, all depend on having the phase velocity in the guiding region lower than that of a surface-wave on the surface away from the guiding region.

A layer, deposited on the substrate, of a material having characteristic acoustic velocities lower than those of the substrate will produce a composite substrate for which the surface-wave velocity is lower than that of the substrate itself. Thus in the left-hand geometry of Fig. 11.2a the phase velocity in the region covered by the slower material will be less than the velocity on the free surface, with the result that the surface-wave energy will be restrained to the region of the plated strip and will be guided by the strip. A typical cross-sectional distribution of the vertical component of displacement is shown below the guide. The thicker the guide for a

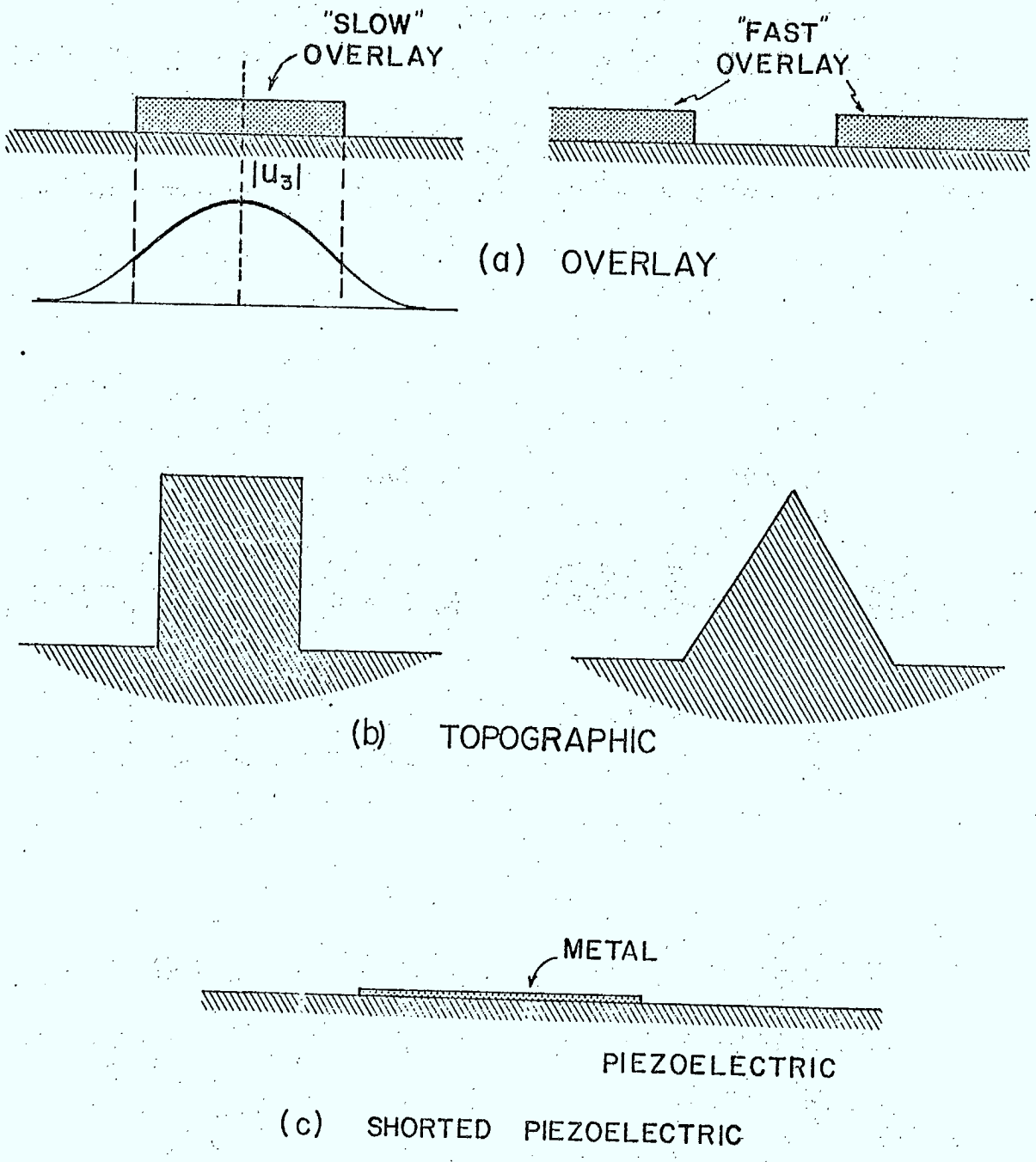


Fig. 11.2 Cross-sectional sketches for three types of acoustic waveguides .

given combination of materials the larger the fraction of the energy that is carried in the overlaid strip and the stronger is the guiding action. While this form of guide is easy to fabricate if the overlay is a sputtered metal, it tends to be dispersive and have propagation losses much higher than with the free substrate.

The geometry on the right of Fig. 11.2a also produces guiding if the overlaid material has characteristic velocities higher than those of the substrate. Here less of the surface-wave energy is generally in the overlaid material so that the losses are lower than for the strip type, also the dispersion is less for a given layer thickness, however, the guiding action is inherently weak.

The topographic guides of Fig. 11.2b will guide a perturbed Rayleigh mode but the guiding is weak unless the ridge is high. For high ridges another mode becomes much more attractive. In this antisymmetric mode most of the energy is contained in the ridge and the displacement is dominantly a side-to-side cantilever motion of the top of the ridge. Such a mode has very small dispersion and has the low losses characteristic of the substrate crystal itself. The open questions concerning the importance of such a guide relate to the convenience of manufacture to system specifications and the ease of excitation at high frequencies.

It has been noted that plating the surface of a piezoelectric with a thin metal layer reduces the velocity of a surface wave on that piezoelectric. Thus in the geometry of Fig. 11.2c the phase velocity along the plated region whose cross-section is shown is less than on the free piezoelectric substrate and a surface-wave propagating in the substrate will be guided by the plated strip (29). Since the fractional change

of velocity with shorting, $\Delta v / v$ or $K^2 / 2$, is small, the guiding here is very weak, however it is convenient to fabricate, has low dispersion and low loss because there is little mechanical energy in the metal. Such guiding would not be useful for waveguide-type components but it has been used to compensate for the diffraction spreading of the acoustic beam which would otherwise occur in a long delay line.

11.3 Long Delays and Reflections

To obtain long delays in acoustic delay lines using bulk waves, the wave can be reflected back and forth between plane bounding surfaces, there is little loss or mode conversion at each reflection, the reflection coefficient is near unity and independent of frequency. A similar broad-band, lumped reflector does not exist for surface-waves. For example, if a surface-wave is incident normally onto a right-angled edge of an isotropic substrate, the amplitude reflection coefficient is about 0.36, the transmission coefficient for the surface-wave down the vertical surface is 0.64, and most important 45% of the energy is converted to bulk modes which radiate from the edge into the substrate (66).

If delays longer than obtainable with conveniently available crystal lengths are required some method of using multiple tracks for the surface-wave beam on the substrate must be found. A meander line constructed from a waveguide would be satisfactory if the strong-guiding waveguides of the previous section were less dispersive and less lossy. One technique which has been used with success involves using top and bottom surfaces of the crystal and rounding the edges as shown in Fig. 11.3 (77), (G4, p.61) so that the surface-wave beam arriving at the edge sees a very

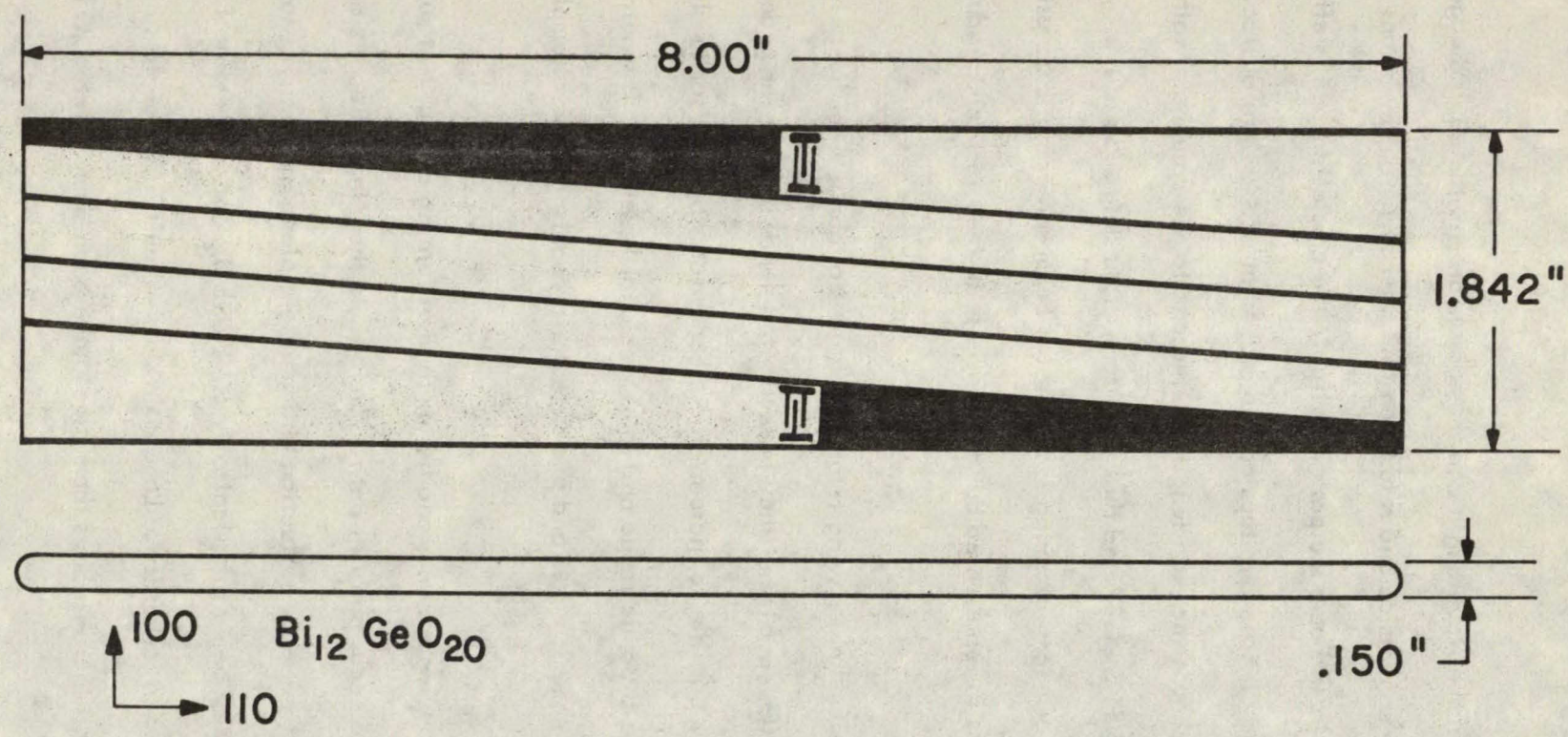


Fig. 11.3 Schematic and relevant parameters of a proposed 1.0 ms, $\Delta f = 60$ MHz, $f_o = 83$ MHz, delay line (77).

gentle curvature in terms of the wavelength and thus passes around the curved edges without reflection or mode conversion and returns on the other surface. By means of a slight inclination of the input beam the path is helical, the crystal surface is efficiently used, the total delay is long and taps can be placed at will. Here diffraction losses become important because the path lengths correspond to the Fraunhofer rather than the Fresnel regions of Fig.11.1, and thus some form of guiding such as the $\Delta v/v$ type of Fig.11.2c or beam reforming is required. Beam spreading is undesirable not only because of the loss so represented but also because of the cross talk produced between adjacent tracks.

Wrap-around delay lines having total time delays of about 1 millisec have been constructed on BGO (G4,p.61) but they have limited bandwidths of the order of 5 to 10 MHz. It should be possible to increase these bandwidths by a factor of 10 while maintaining tolerable dynamic range and spurious signal levels, but it would require an appreciable development effort and in particular probably the use of amplification at intervals along the track.

The multistrip coupler provides another method of turning an incident surface-wave beam back onto an adjacent parallel path. Such track changing reflectors are broad band, reform the beam at each reflection and maintain planar geometry; however they have the disadvantage of being practical only on $Li\,Nb\,O_3$ and at present of having appreciable ohmic loss at each reflection. Long delays using multistrip couplers have not yet been reported though the reflectors themselves have been demonstrated (G10,p.60).

XII. SURFACE-WAVE AMPLIFIERS

Although surface-wave amplifiers have as yet not proved to be viable alternatives to transistor amplifiers, there are a number of surface-wave device applications where it might be advantageous to incorporate as an intergral part of the device some surface-wave amplification. The most important application would be in obtaining loss compensation in long delay lines of the wrap-around type and in obtaining triple-transit echo reduction in surface-wave devices by virtue of the inherent isolation that results from differences in the forward and reverse gain of the amplifier.

Amplification of surface-waves is a result of the interaction of the electric fields accompanying the surface acoustic wave ona piezoelectric surface with drifting carriers either in an adjacent semiconductor or in the piezoelectric itself, if it is also semiconducting. Gain is obtained when the drift velocity v_D exceeds the surface-wave velocity v. This type of amplification has the advantage over external methods in that distributed gain compensates for distributed loss, high gain and bandwidth is achieved without interstage coupling networks, and isolation is obtained which helps eliminate echoes and other undesirable spurious signals.

The gain which is possible using surface-wave amplifiers depends on material parameters and device geometry. Many configurations have been proposed and analytical and experimental studies have been carried out to demonstrate their usefulness at VHF and UHF. The performance of these amplifiers can be analyzed (in principle) by solving the appropriate electromechanical field equations for the piezo-

electric in conjunction with the transport equations for the carriers in the semiconductor subject to the boundary conditions imposed by the specific device geometry. Such an approach leads to involved and expensive computations which do not give much insight into the effect of physical parameters and device geometry and a number of perturbation theories have been developed which focus attention on some key physical parameters which have proven to be useful guidelines for the detailed numerical calculations needed for the design of a specific amplifier.

In the following sections, the different methods used to amplify surface-waves are described and the advantages of the various configurations summarized. The key design parameters are discussed and typical experimental and theoretical results given for each amplifier type.

12.1 Amplifier Configurations

Surface-wave amplifiers can be classified in accordance with the methods and the materials used to obtain acoustoelectric interactions into the following four categories.

(a) A combined-medium amplifier, as its name implies, utilizes a piezoelectric semiconductor such as CdS or GaAs as both the propagating and the active medium. A typical configuration showing transmitting and receiving transducers and the carrier drift field region is given in Fig. 12.1.

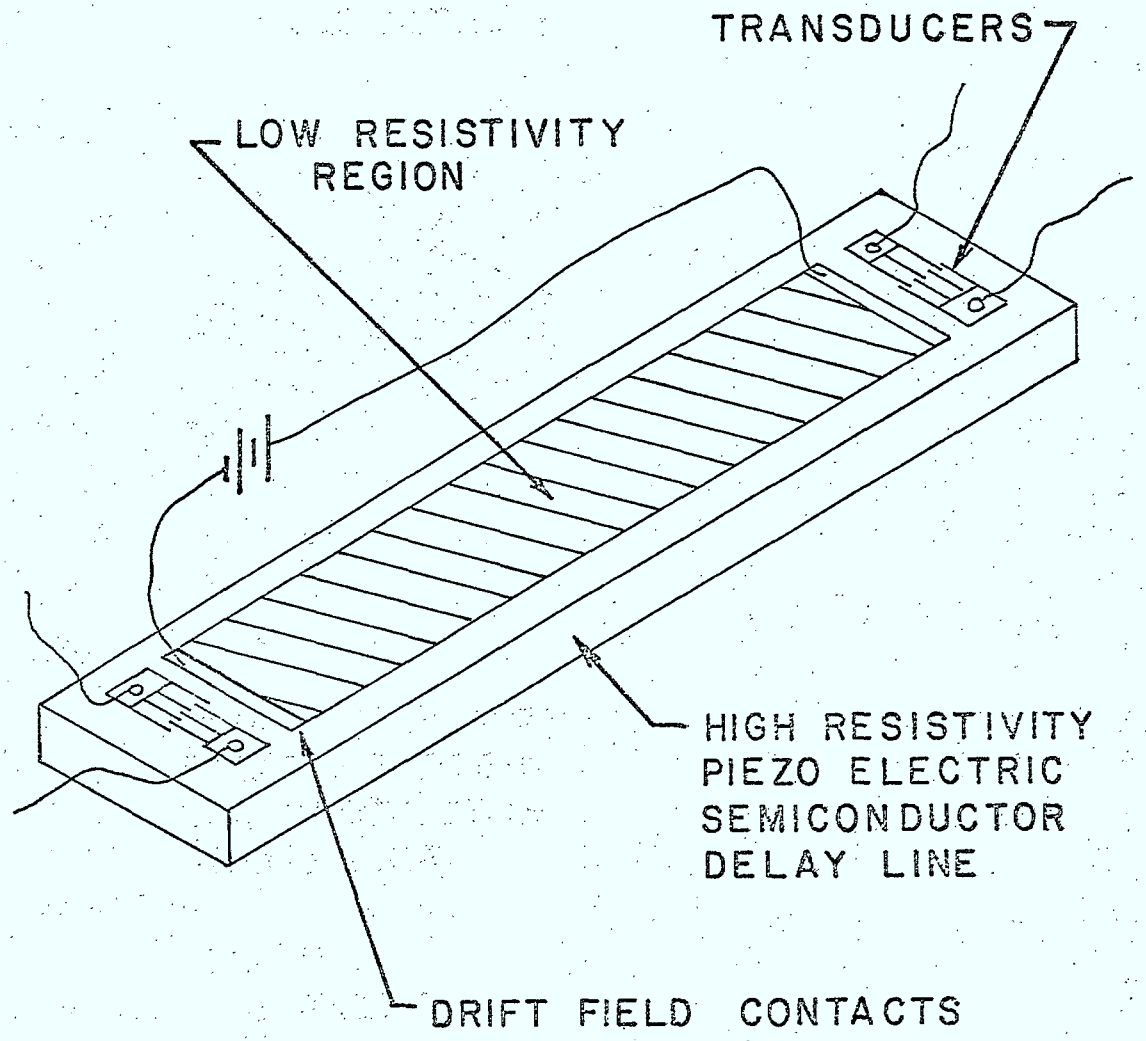


Fig. 12.1 Combined-medium surface-wave delay line amplifier (48).

(b) A separated - medium amplifier uses a piezoelectric insulator as the propagating medium and a biased semiconductor as the active medium arranged to be adjacent but separated by a small air gap in order to avoid mechanical loading of the surface-wave and hence keep the device dispersion very low. Such a structure is shown in Fig. 12.2, the acoustoelectric coupling exists because the electric fields which accompany the surface-wave on the piezoelectric extend across the air gap into the semiconductor. The effect of air gap on coupling is illustrated in Fig. 12.3 which shows the gain vs frequency for a separated - medium amplifier for two values of air gap width (see Section 12.3 below).

(c) A monolithic amplifier uses a semiconductor film on a piezoelectric substrate (or vice versa) as part of a layered structure. For example, an amplifier using an In Sb film on a Li Nb 0_3 substrate is shown in Fig. 12.4 .

(d) A strip-coupled amplifier consists of a semiconductor region in which drifting carriers are coupled via a multistrip coupler to an acoustic surface-wave propagating on a piezoelectric surface. A number of multistrip amplifier configurations have been proposed, two of which are shown in Figs. 12.5 and 12.6. It should be

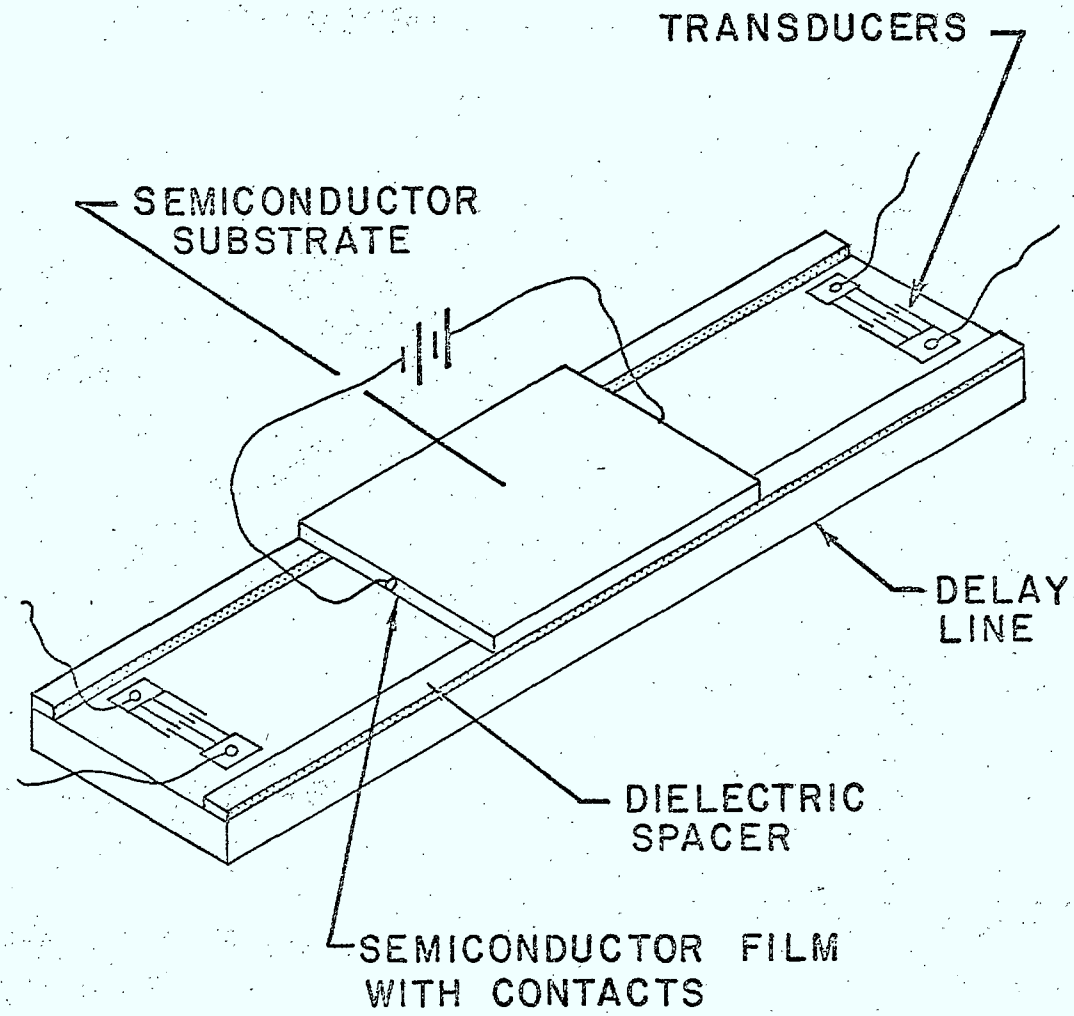


Fig. 12.2 Separated-medium surface-wave delay line amplifier (48).

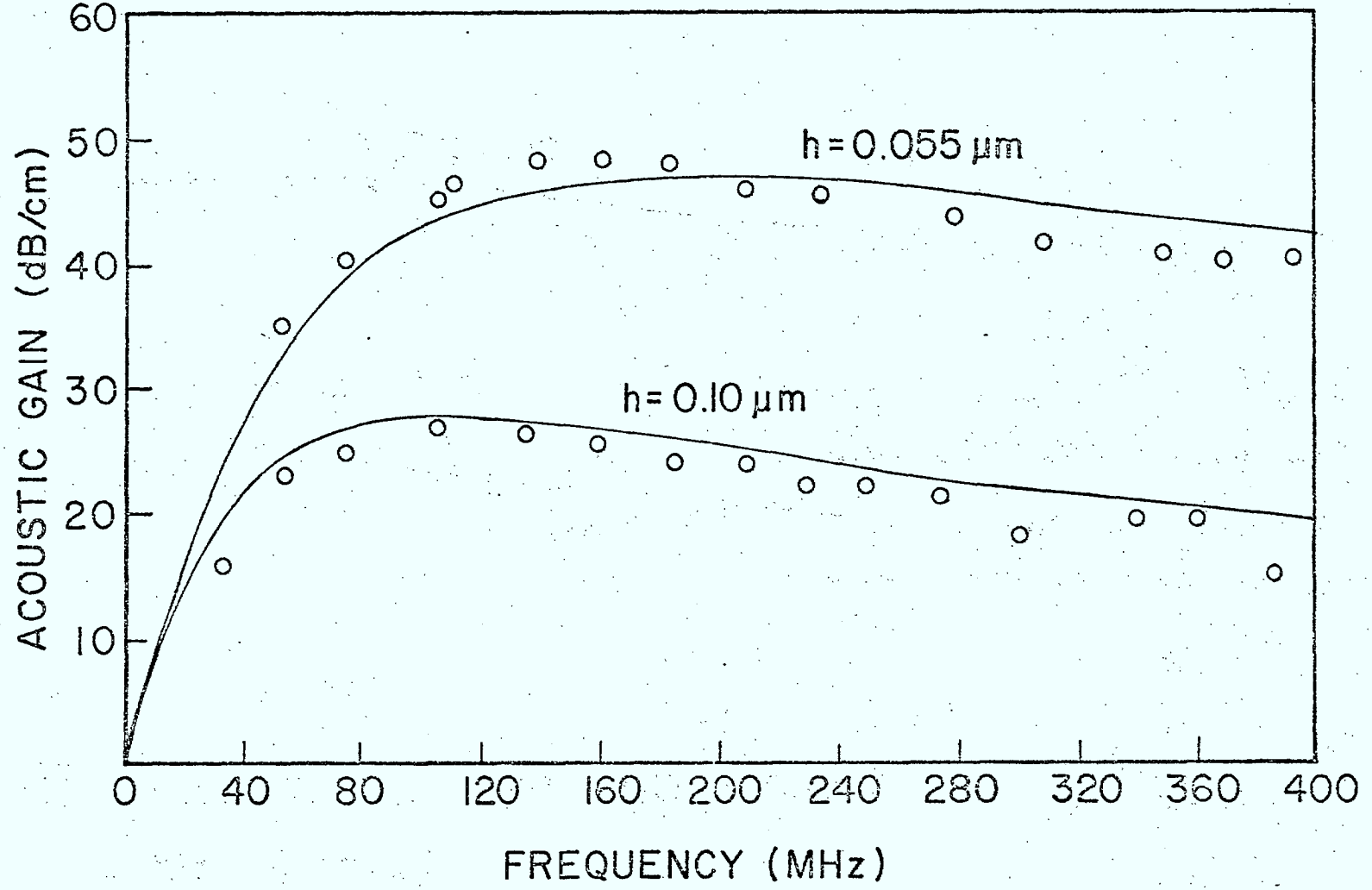


Fig. 12.3 Gain versus frequency for Li $Nb\,0_3$ and silicon-on-sapphire separated-medium amplifier (43).

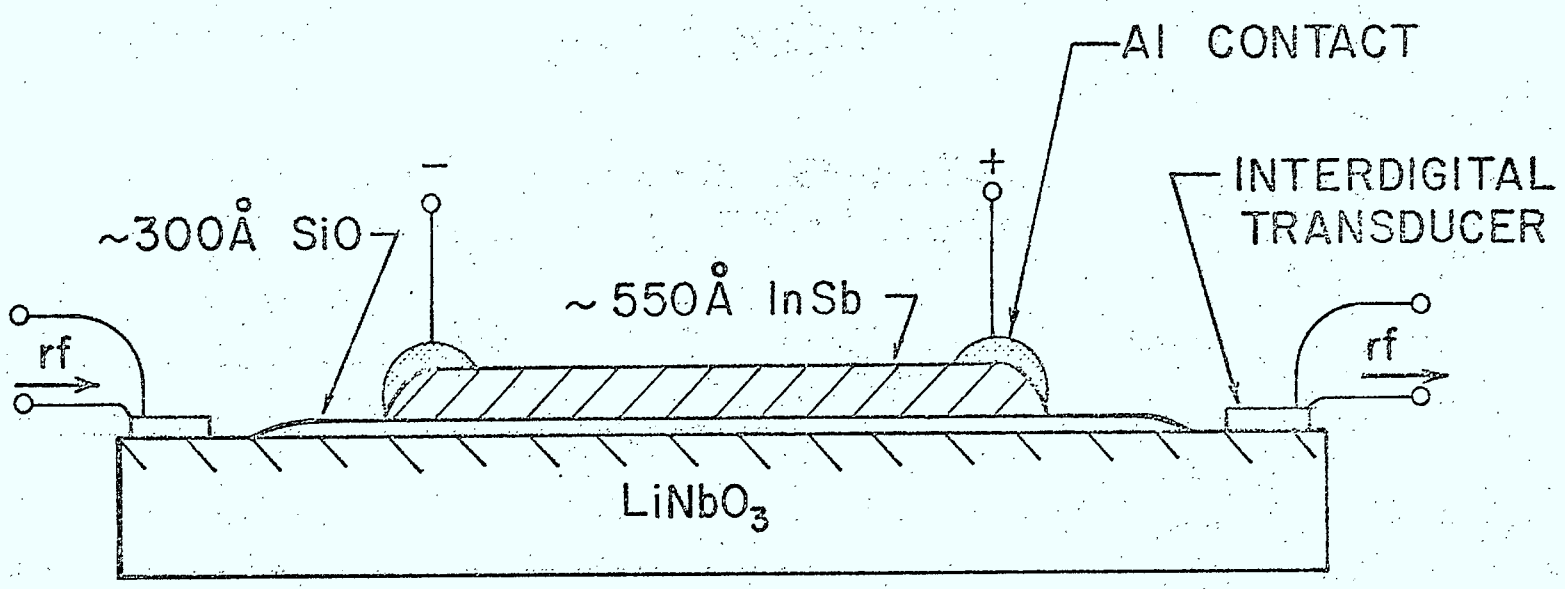


Fig. 12.4 Monolithic amplifier configuration (13).

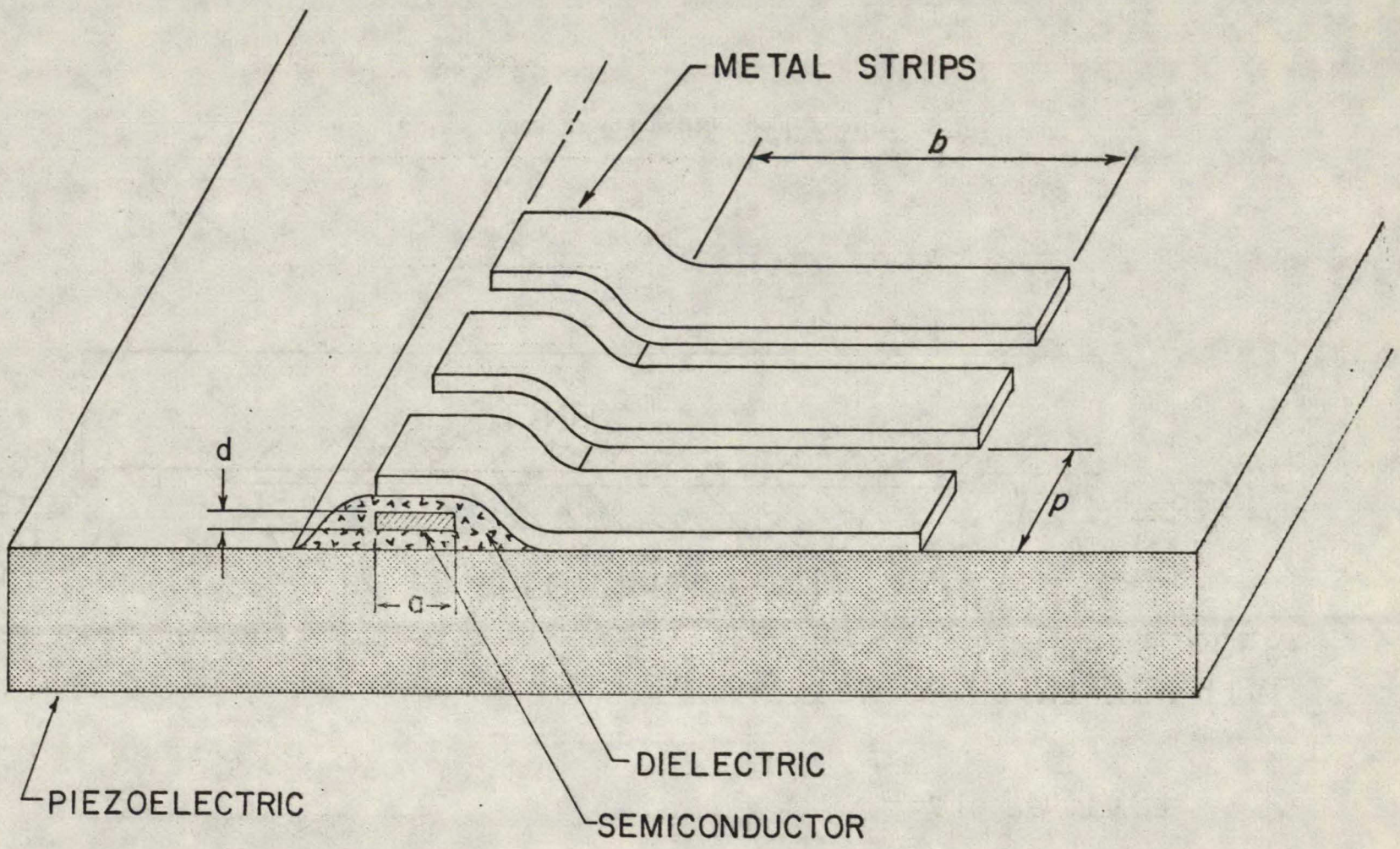


Fig. 12.5 Strip-coupled surface-wave amplifier (39) .

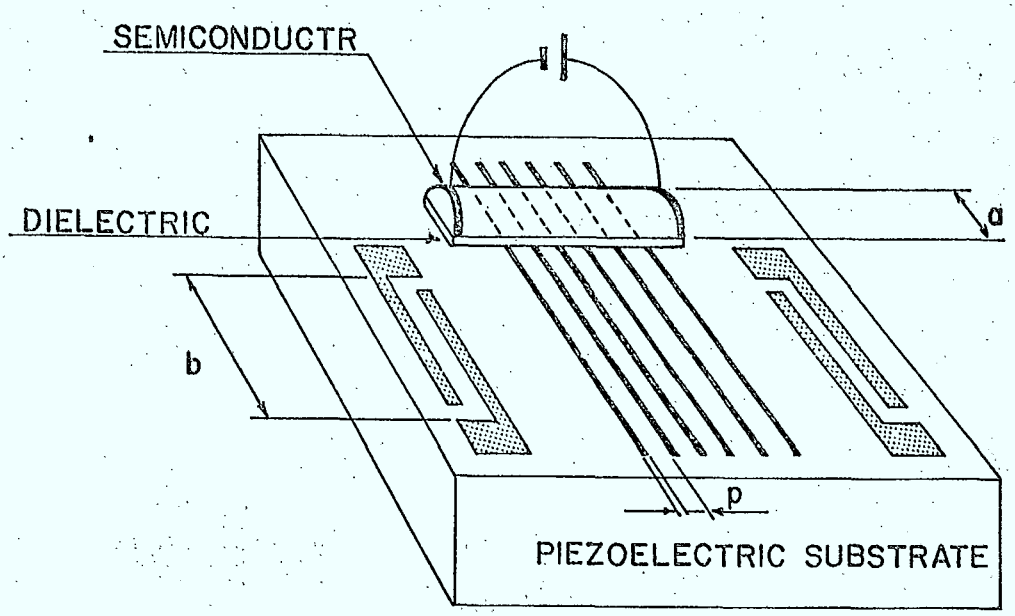


Fig. 12.6 Strip-coupled amplifier with semiconductor overlay (G5, p. 171).

noted that to date no amplifier of this type has been successfully demonstrated experimentally.

At this point it is useful to contrast these different configurations from the point of view of material requirements, fabrication techniques and salient operating features.

12.2 Combined Medium Amplifier

This configuration, Fig. 12.1, is clearly the simplest amplifier to construct, requiring only a polished surface on a piezoelectric semiconductor on which transducer and drift field electrodes are fabricated. Unfortunately materials are apparently not simultaneously good piezoelectrics and high mobility semiconductors so that even Zn O, CdS and GaAs, which are the best known piezo-semiconductors do not have the optimum parameter combinations. The first two are good piezoelectrics but have a rather low mobility, the last named has a high mobility but a rather low electromechanical coupling constant.

Cadmium sulfide which is also a photoconductor was the first material used in surface-wave amplifier experiments (103, 105). The experimental results obtained compared well with theoretical predictions using a bulk-wave type of analysis. More recently (54) in experiments using transverse surface-waves on GaAs electronic gains were measured ranging from 35 to 50 dB / cm at 205 MHz and from 75 to 80 dB / cm at 615 MHz. These results are in good agreement with theoretical predictions over a wide range of resistivities (changed by heating)

and drift fields. A serious disadvantage of the combined medium structure is that the size of the semiconductor part is the full size of the device so that a lot of power is supplied by the drift field which tends to cause serious heating problems. Most, if not all, combined medium experiments have had to use low repetition rate pulsed drift fields.

12.3 Separated Medium Amplifier

This configuration shown in Fig. 12.2 has the obvious advantage that the piezoelectric and semiconductor materials can be chosen independently for an overall optimum performance. Furthermore the semiconducting region geometry can be chosen such that the power from the drift field is modest and thus heat sinking is not a serious problem. Separated medium amplifiers have little dispersion since there is no mechanical loading of the piezoelectric because of the existence of the air gap. The air gap requirement, however, results in severe demands on the fabrication techniques from the point of view of spacing tolerances and optical finish of the surfaces. Dielectric spacer rails, and randomly distributed small area posts obtained by sputter machining have been used to control air gap spacing (G6, p.142).

A great deal of theoretical and experimental work has been carried out on this structure (43, 47), for example, a Silicon - Li Nb 0_3 configuration was used (48) to demonstrate its feasibility at UHF frequencies. In order to reduce the drift voltages needed and to obtain better heat sinking, the silicon was epitaxially grown in a segmented structure on a sapphire substrate. The sapphire acts as a heat sink and the segments are biased in parallel by a modest voltage. The silicon side is placed about 560 A° from the Li Nb 0_3 by means of dielectric spacers as shown in Fig. 12.2.

This particular structure yielded 40 dB/cm of gain at 108 MHz with a bias of 185 volts and about 3 dBm of output power is available before the amplifier saturates. Calculations (40) indicate that noise figures as low as 5 dB are achievable and that carrier trapping is the major source of noise degradation. Figure 12.7 shows experimental and theoretical noise figure results at 106 MHz as a function of electronic gain. The best fit to the experimental data occurs when 82% of the carriers are untrapped ($f = 0.82$).

The theory for the separated medium amplifier is rather complicated but a perturbation technique has been used (48) to derive a formula for the attenuation

$$\alpha = \mathrm{Im}\left\{\beta\left(\frac{\Delta v}{v}\right)\left(\frac{1-\tanh\beta h}{1+\frac{\epsilon_s}{\epsilon_o}\tanh\beta h}\right)\left(\frac{1+i Z_s}{1-i g Z_s}\right)\right\} \qquad (12.1)$$

where β is the propagation constant, h is the air gap, and Z_s is a dimensionless impedance for the semiconductor defined by $Z_s = \epsilon_o E_1/D_3$ where E_1 is the tangential field parallel to the propagation direction and D_3 is the electric displacement normal to the interfaces, both evaluated at the semiconductor surface, and g is defined by

$$g = \frac{\epsilon_p + \epsilon_o \tanh\beta h}{\epsilon_o + \epsilon_s \tanh\beta h} ,$$

with $$\epsilon_s = \sqrt{\epsilon_{33}\epsilon_{11} - \epsilon_{13}^2} .$$

Note the role of $\Delta v / v$ in the gain equation. Essentially the same results are

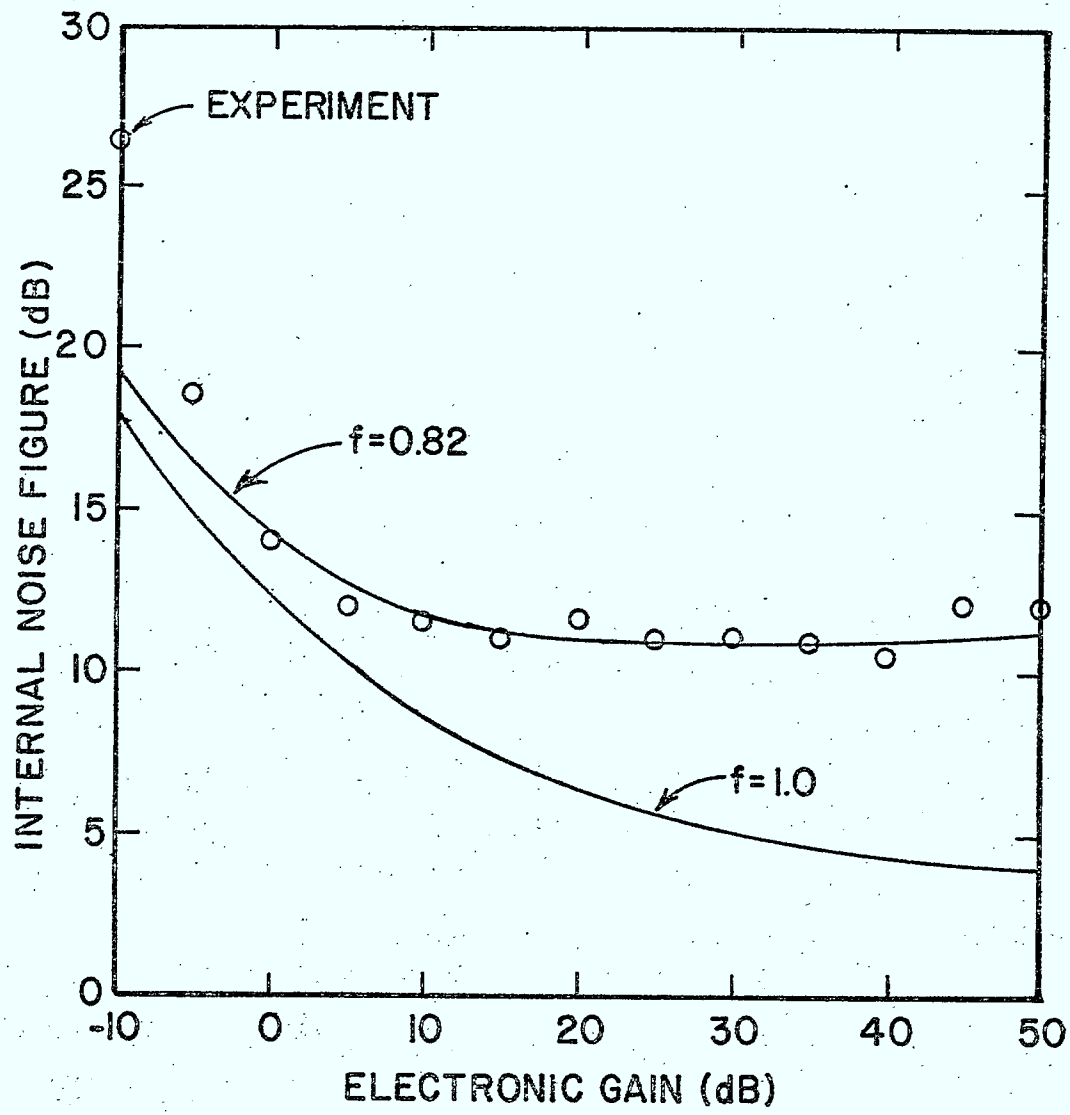


Fig. 12.7 Theoretical and experimental noise figure versus electric gain for a separated-medium amplifier (40).

obtained by other perturbation formulas (33), (G6, p.275). Since most amplifiers of this type would operate at frequencies where diffusion effects are negligible Eq. 12.1 can be written

$$\alpha = \beta (\Delta v/v) \cdot \frac{1 - \tanh \beta h}{1 + \frac{\epsilon_s}{\epsilon_o} \tanh \beta h} \cdot \frac{\eta R}{\eta^2 + R^2} \qquad (12.2)$$

where $\eta = 1 - v_D/v$ with v_D the carrier drift velocity, and

$$R = \frac{\sigma d \tanh \beta d}{v \epsilon (h) \beta d} \cdot \qquad (12.3)$$

σd is the sheet conductivity of the semiconductor and $\epsilon (h)$ is an effective permittivity which depends on ϵ_s and the air gap height and is given by

$$\epsilon (h) = (\epsilon_s + \epsilon_o)(1 + \tanh \beta h) / (1 + \frac{\epsilon_s}{\epsilon_o} \tanh \beta h) . \qquad (12.4)$$

The maximum gain occurs when $\eta = R$ at a drift velocity $v_D = 1 + R$ and is

$$\alpha_{MAX} = \frac{\beta}{2} \cdot \frac{\Delta v}{v} \cdot \frac{1 - \tanh \beta h}{1 + \frac{\epsilon_s}{\epsilon_o} \tanh \beta h} \cdot \qquad (12.5)$$

From Eqs. 12.1 to 12.5 it can be seen that the maximum gain is proportional to $\frac{\Delta v}{v}$ the key electromechanical quantity discussed in Chapter X, multiplied by a rather complicated function of frequency.

Many separated-medium amplifiers have been fabricated to operate at frequencies ranging from 40 MHz to 1 GHz. To illustrate their capabilities in Fig.12.3 a plot of gain vs frequency for silicon-on-sapphire semiconductor region and a

YZ Li Nb 0_3 substrate has been given for two values of air gap. The solid curve is theoretically calculated (43), the data points are experimental results, the critical amplifier parameters are:

$$v_D / v = 3 ,$$

$$\sigma d = 11.5 \times 10^{-6} (\Omega/\square)^{-1} ,$$

$$d = 1 \mu m .$$

From the curves, or by examining Eq. 12.5 it can be seen that at low frequencies the gain increases linearly and is independent of air gap; as the frequency is increased the gain is almost independent of frequency but inversely proportional to h .

A detailed analysis of the amplifier performance reveals that in order to keep the device power dissipation within reasonable limits and still yield usable gain the sheet conductivities must be kept in the range 10^{-6} to $10^{-5} (\Omega/\square)^{-1}$.

The main disadvantage of this structure lies in the air gap tolerances which must be maintained especially at high frequencies and it is doubtful if such a configuration will prove directly viable in any system application. However, similar performance is obtained by supporting the semiconductor on many small randomly located "posts" produced on the substrate surface by shallow etching into the polished surface (G6, p.142) .

12.4 Monolithic Structures

This is essentially a separated-medium structure with zero air gap and thus also avoids one of the main limitations of the prototype separated-medium geometry. The

choice and compatibility of materials and fabrication techniques introduces other constraints, quite apart from the extra dispersion introduced by the layered structure. The main advantage of a monolithic amplifier lies in the possibility of integrating its fabrication as part of a microelectronic assembly line.

The first successful experiments used a Cd Se film vacuum deposited onto Li $Nb O_3$ (25) and gave gains of 10 dB per mm (100 dB / cm) at 170 MHz. Similar results were obtained independently (34) for the same structure and it was shown that the large discrepancies between experimental results and theory were due to a depletion layer and carrier traps at the Li $Nb O_3$ semiconductor interface. The most successful structure tested (13) consisted of a 550 $\mathring{A}$ In Sb film flash evaporated on a Li $Nb O_3$ substrate on which a thin insulating (300 $\mathring{A}$) layer of $Si O_2$ was first deposited, this configuration is shown in Fig. 12.4. The gain vs field for this amplifier is given in Fig. 12.8 along with a theoretical curve given by the Kino-Reeder normal mode theory (43). The amplifier gave 63 dB/cm of electronic gain at 660 MHz under pulsed conditions, implying a net terminal gain of 20 dB since the delay line loss was a rather large 43 dB. The linearity characteristic for the amplifier is given in Fig. 12.9 from which is seen that the output saturates at about 20 dBm .

The theoretical analysis of monolithic amplifiers containing many layers is beyond the scope of this report but it should be mentioned that the structure consisting of Zn O on high resistivity silicon having a thin doped or ion implanted channel is potentially capable of gains as high as 600 dB/cm at UHF frequencies (G5,p.440), (G6,p.275).

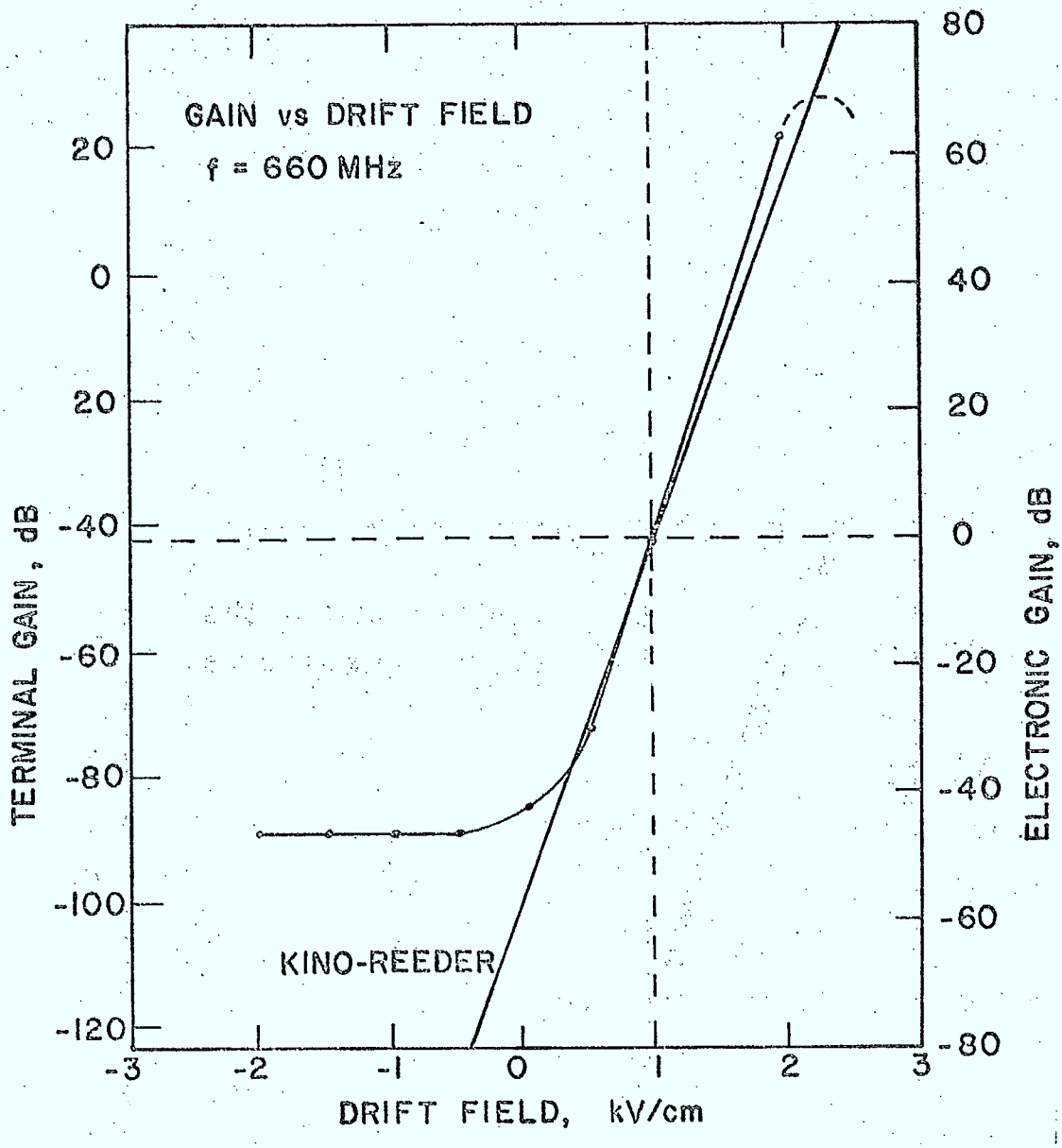


Fig. 12.8 Gain versus drift field for monolithic amplifier of Fig. 12.4 (13).

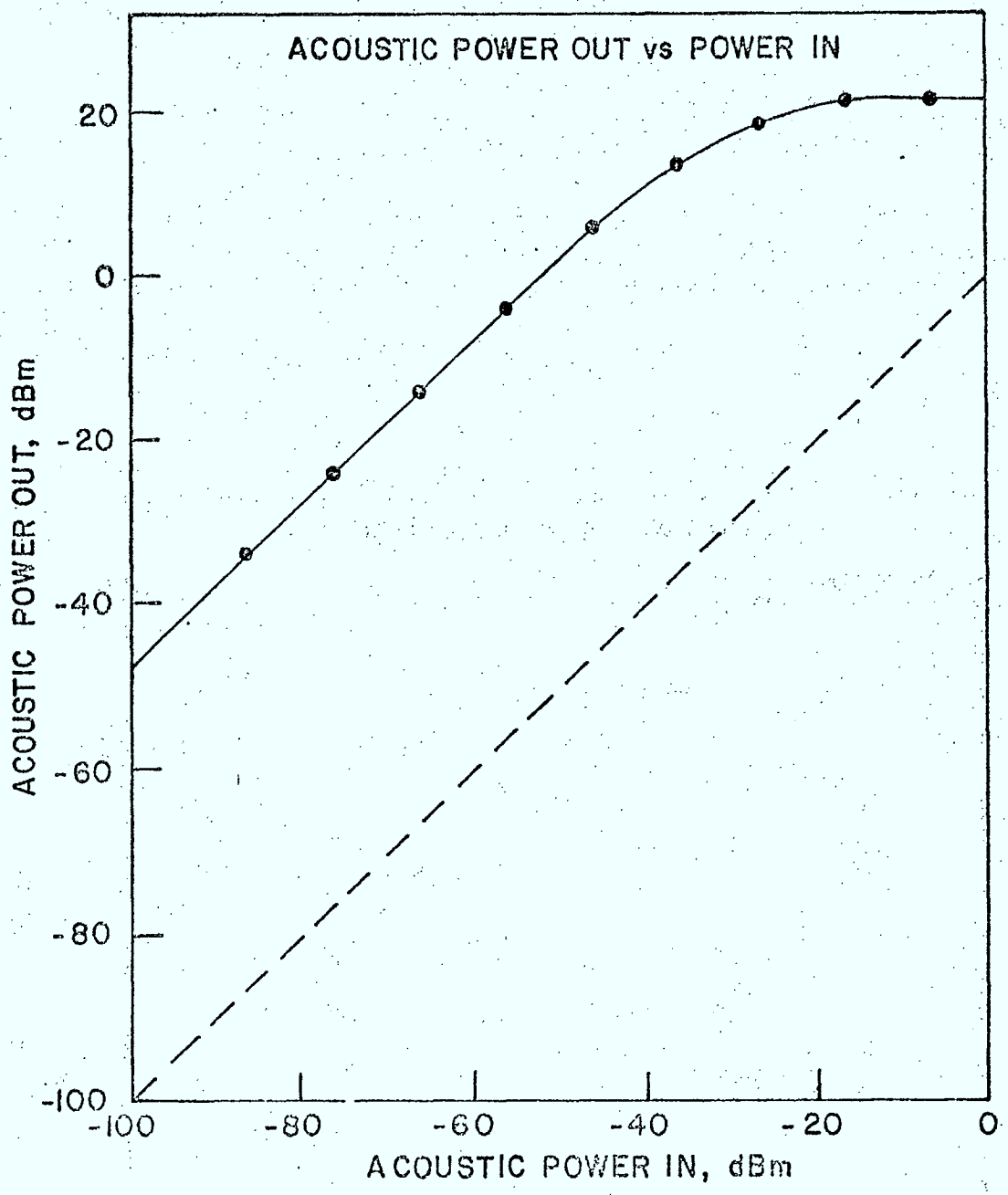


Fig. 12.9 Saturation characteristic for monolithic amplifier of Fig. 12.4 (13).

12.5 Strip Coupled Amplifier

This configuration is by far the most flexible, allowing considerable choice in materials, geometry, and operating conditions. In Fig. 12.5 a schematic drawing of a section of a strip coupled amplifier is given (G5, p.179). The surface-wave is excited by a separate transducer (not shown) of width b, and propagates on the piezoelectric in a direction normal to the strips and parallel to the semiconductor. The strips which are insulated from the semiconductor by a dielectric film in effect capacitively couple the electric field of the wave to the carriers in the semiconductor. When a voltage is applied to the semiconductor film to produce a drift velocity in the direction of propagation which exceeds the wave velocity, amplification is obtained. Fig. 12.6 shows the entire amplifier structure for a slightly different geometry in which the semiconductor is deposited onto the strips with a thin insulating dielectric layer between the strips and the semiconductor. The ability to have a semiconducting strip which is narrower than the acoustic beam width makes it possible to reduce the power dissipation of the resulting amplifier without sacrificing gain (2). Note also that the periodicity and the metallization ratios at the semiconductor can differ from those on the acoustic path.

Since this structure is relatively new, and no net gain has been measured to date, only theoretical evaluations of the strip coupled amplifier have been carried out. Using an FET theory it has been shown (39) that the attenuation for a device having many strips per wavelength can be written

$$\alpha = \beta \left(\frac{\Delta v}{v}\right) \frac{C}{C_b} \frac{\eta R}{\eta^2 + R^2} \qquad (12.6)$$

where C_b is the capacitance of the strips to ground, C is the capacitance of C_b in series with the capacitance between the metal strips and the semiconductor,

$$R = \frac{\sigma d}{v} \frac{\beta a}{C} , \quad \eta = 1 - v_D/v ,$$

and a is the width of the semiconductor and d the height. Equation 12.6 gives the same results as the space harmonic theory (7), (G5,p.171) under the same assumptions of many strips per wavelength. The predicted gain (G5,p.171) of such an amplifier is shown in Fig.12.10 as a function of relative drift velocity for a strip width to period ratio of 0.5, for the entire strip structure and a conductivity chosen such that the dielectric relaxation frequency equals the operating frequency. The running parameter γ is the strip period to wavelength ratio p/λ and the gain is given in dB per wavelength normalized to the surface-wave coupling constant K^2. The following remarks can be made as a result of such calculations:

1. The maximum gain decreases as γ goes from zero (continuous coupling) to 0.5.
2. Gain is possible for reverse fields at velocities lower than the acoustic velocity.
3. Gain curves become symmetrical about the vertical axis for $p/\lambda = 0.5$.
4. For $p/\lambda > 0.4$ gain is not possible at high conductivities with reasonable drift velocities.

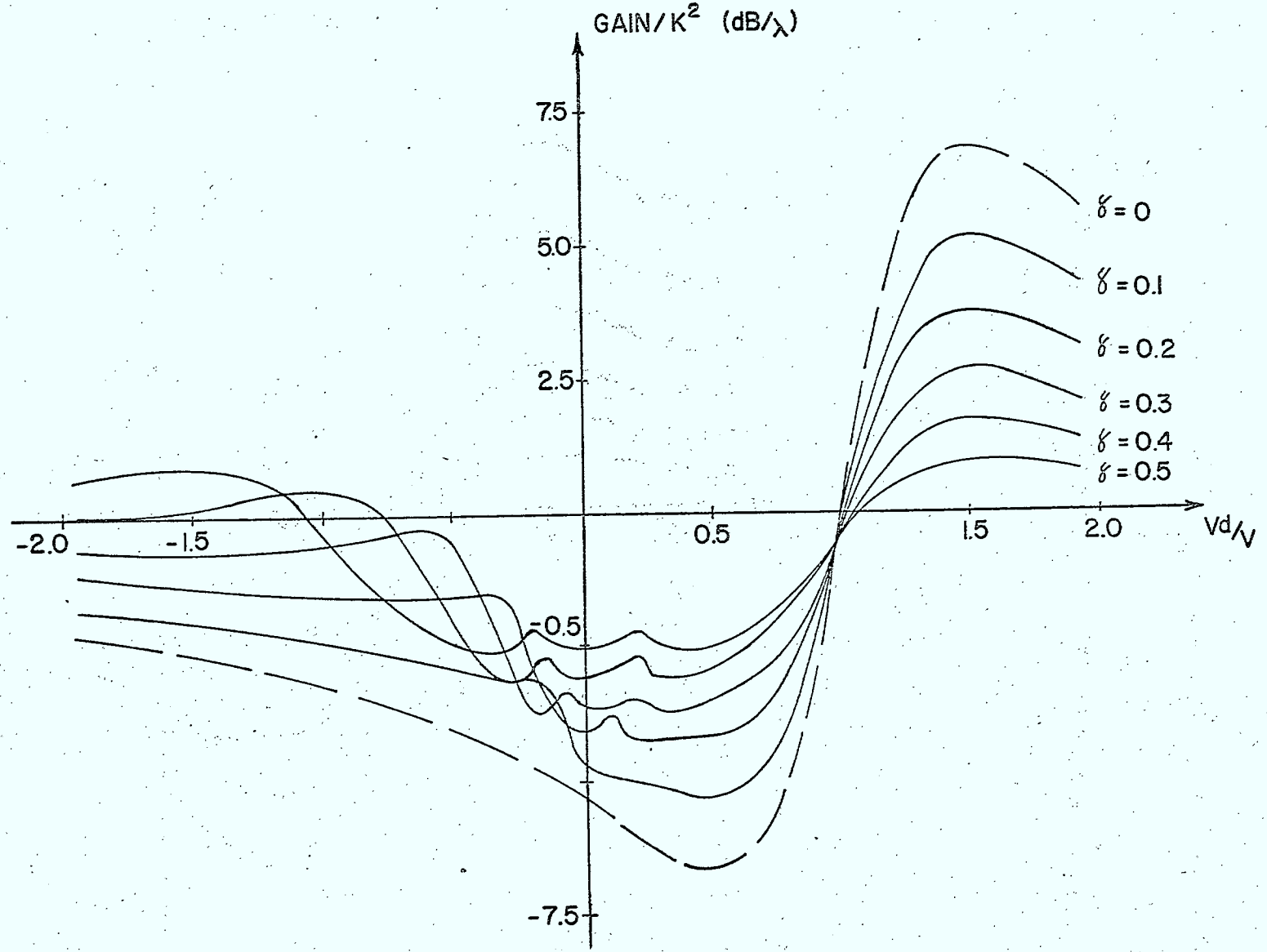


Fig. 12.10 Theoretical gain versus relative drift velocity for strip-coupled amplifier of Fig. 12.6 . Running parameter γ is strip to wavelength ratio p/λ (G5, p.171) .

In view of the lack of experimental verification for strip coupled amplifiers and the severe technological problems encountered in fabricating them, it is unlikely that such amplifiers will be anything but experimental devices in the near future (G10).

XIII. SURFACE-WAVE CONVOLVERS

13.1 Principle of Operation

In the preceding chapters many different surface-wave signal-processing devices such as bandpass filters, tapped delay-lines, and dispersive delay lines have been discussed. With nearly all these devices the desired transfer function or impulse response is achieved by properly choosing the width, spacing, and overlap of an interdigitated configuration. The signal processing operations obtained are fixed or time invariant in that a given filter has a fixed impulse response for a given transducer pattern and the system output is the convolution of the signal with the filter impulse response so that such a filter will be matched to only one signal. It is desirable to have available filters whose transfer functions are electronically variable or equivalently devices to perform the real time correlation or convolution of signals. Such devices can be realized through nonlinear interactions in surface-wave delay lines and are capable of performing many of the operations done by linear filters, but with electronic programmability.

These signal-processing devices employ the parametric interactions of surface-waves passing in opposite directions along a delay line and are capable of giving the real-time convolution of two modulated signals. Since one signal acts as the reference for the other, a large number of electronically variable signal-processing functions are possible such as recognition of digital codes, compression of FM chirps, and the generation of the Fourier transform of signals in real-time (71).

As an indication of the structures involved and as motivation the simplest surface-wave convolver is shown in Fig. 13.1. This device consists of a piezoelectric substrate with one transducer at each end corresponding to the two parts labelled 1 and 2 and a third port formed by two metal electrodes deposited on the top and bottom surfaces of the crystal and midway between the two transducers. When two oppositely travelling RF modulated signals of the same frequency ω with envelope modulation $F(t)$ and $G(t)$ are launched by the transducers the nonlinear effects give rise to a product term which results in an output signal at twice the frequency ω and with envelope modulation $C(t)$ of the form

$$C(t) = -B v \int_{-\infty}^{\infty} F(\tau)\, G(2t-\tau)\, d\tau \quad , \tag{13.1}$$

where B is a nonlinear coefficient for the parametric interaction and v the velocity. The integration limits are infinite when F and G are of smaller duration than the transit time under the pick up plate. This output is the convolution of F and G time compressed by a factor of two and is detected across the two centre plates at port 3. Correlation is obtained by applying an RF signal at frequency ω with envelope $F(t)$ to port 1 and a pump field at frequency $2 \times \omega$ with modulation $G(t)$ at port 3. Returning to port 1 is a signal at a frequency ω with envelope modulation

$$C(t) = -\mathcal{L} v \int_{-\infty}^{\infty} G(\tau)\, F(2\tau-t)\, d\tau \quad , \tag{13.2}$$

where $\mathcal{L}$ is a nonlinear coefficient. This signal is the correlation of F and G time expanded by a factor of 2.

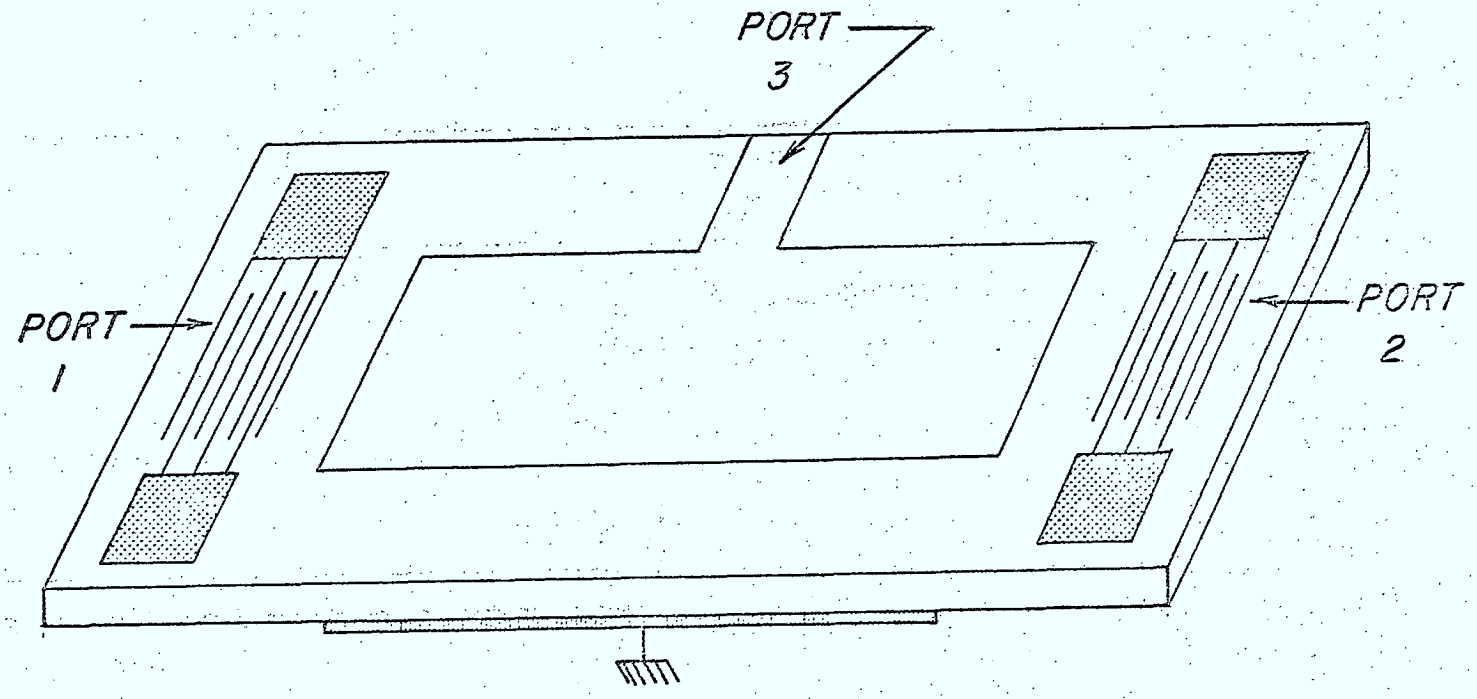


Fig. 13.1 Schematic representation of a surface-wave degenerate convolver.

In principle any processing requiring the convolution or correlation operations can be processed in real-time with a single device, this is the programmable feature of a parametric processor which distinguishes it from fixed linear filters.

In this chapter the sources of the nonlinearity giving rise to the parametric interactions will be briefly discussed and then the remainder of the chapter will deal in some detail with the main types of convolvers which have been evaluated together with performance characteristics where appropriate.

13.2 Surface-Wave Nonlinearity

The expectation that strong nonlinear effects would occur with surface-waves has been verified experimentally in a number of laboratories. Harmonic generation has been detected on Quartz (49, 53), Lithium Niobate (50, 80), (G6, p.268) and $Bi_{12}\,Ge\,0_{20}$ (1). In these experiments a laser probe technique is used to measure second and higher harmonics in the surface-wave and also to determine at what acoustic power levels saturation of the fundamental frequency occurs. The results given in (80) show clearly the power handling limitations in surface-wave devices, for example a maximum power of 10 mW / mm at 1 GHz is given as a limit for linear operation on YZ Li Nb 0_3 this limit is inversely proportional to frequency.

The source of nonlinearity is of course the quadratic terms in the constitutive equations or the cubic terms in the energy. To illustrate in one dimension these equations can be written (55)

$$T = C_1 S - e_1 E + \frac{1}{2} C_2 S^2 - \frac{1}{2} d_2 E^2 - f_2 E S$$

$$D = \epsilon_1 E + e_1 S + \frac{1}{2} \epsilon_2 E^2 + d_2 E S + \frac{1}{2} f_2 S^2 ,$$

where the finite strain $S = \frac{\partial u}{\partial x} + \frac{1}{2} (\frac{\partial u}{\partial x})^2$.

The suffix 1 refers to effective linear coefficients and the suffix 2 to effective quadratic coefficients. Thus we can identify d_2 as an electrostrictive coefficient, ϵ_2 as an electro-optic coefficient etc. In fact for anisotropic materials these are all tensor quantities and there are thus many of these coefficients to be identified. For YZ Li Nb 0_3 some of the third order elastic constants have been measured (67) and they are only known for a very limited number of materials (52).

There have been many attempts to analyze the problem of harmonic generation of surface-waves (101), (G6,p.244), most of which use some form of perturbation to first calculate the initial growth rate of the second harmonic, and some success has been obtained (G6,p.268) at modelling harmonic generation using one nonlinear parameter to characterize the surface-wave nonlinearity. An exact or numerical solution to this nonlinear boundary value problem has not yet been obtained.

In addition to inherent material nonlinearity it is possible to obtain strong nonlinear effects by coupling the surface wave electric fields to an adjacent semiconductor wafer in configurations very similar to the separated-medium and strip-coupled amplifiers discussed in Chapter XII, it turns out that the nonlinearity arising from acoustoelectric interactions is so large that semiconductor-coupled convolvers have out-performed all other types.

13.3 Surface-wave Convolution

To understand the basic signal-processing function of a surface-wave convolver consider two acoustic signals $F(t)$ and $G(t)$ launched from opposite ends of a delay line and propagating towards each other with the surface-wave velocity v. The x-axis is chosen parallel to the direction of propagation and the origin placed at the centre of the delay line. The basic surface-wave convolver configuration is shown in Fig. 13.2. Contra-directed surface-wave trains are excited in the delay line having the form $F(t - \frac{x}{v}) \exp j(\omega_1 t - k_1 x)$ and $G(t + \frac{x}{v}) \exp j[\omega_2 t + k_2 x]$, of synchronous frequency ω_1 and ω_2, and opposite wave vectors k_1 and k_2. These waves mix via nonlinearities as they slide through one another to generate an electric polarization at the sum of the two surface wave frequencies $\omega_3 = \omega_1 + \omega_2$ but with a wave vector $k_3 = k_1 - k_2 = \frac{\omega_1 - \omega_2}{v}$. This sum frequency signal is detected by the centre output transducer designed to have a periodicity corresponding to the wave number k_3. Since ω_3 / k_3 is not equal to the phase velocity this product signal is a standing wave and not a travelling wave and the output transducer integrates and detects this product signal to give an output modulation at ω_3 given by

$$C(t) = B \int_{\frac{L}{2}}^{\frac{L}{2}} F\left(t - \frac{x}{v}\right) G\left(t + \frac{x}{v}\right) dx ,$$

where L is the length of the centre pick up transducer and B is a nonlinearity constant for the interaction.

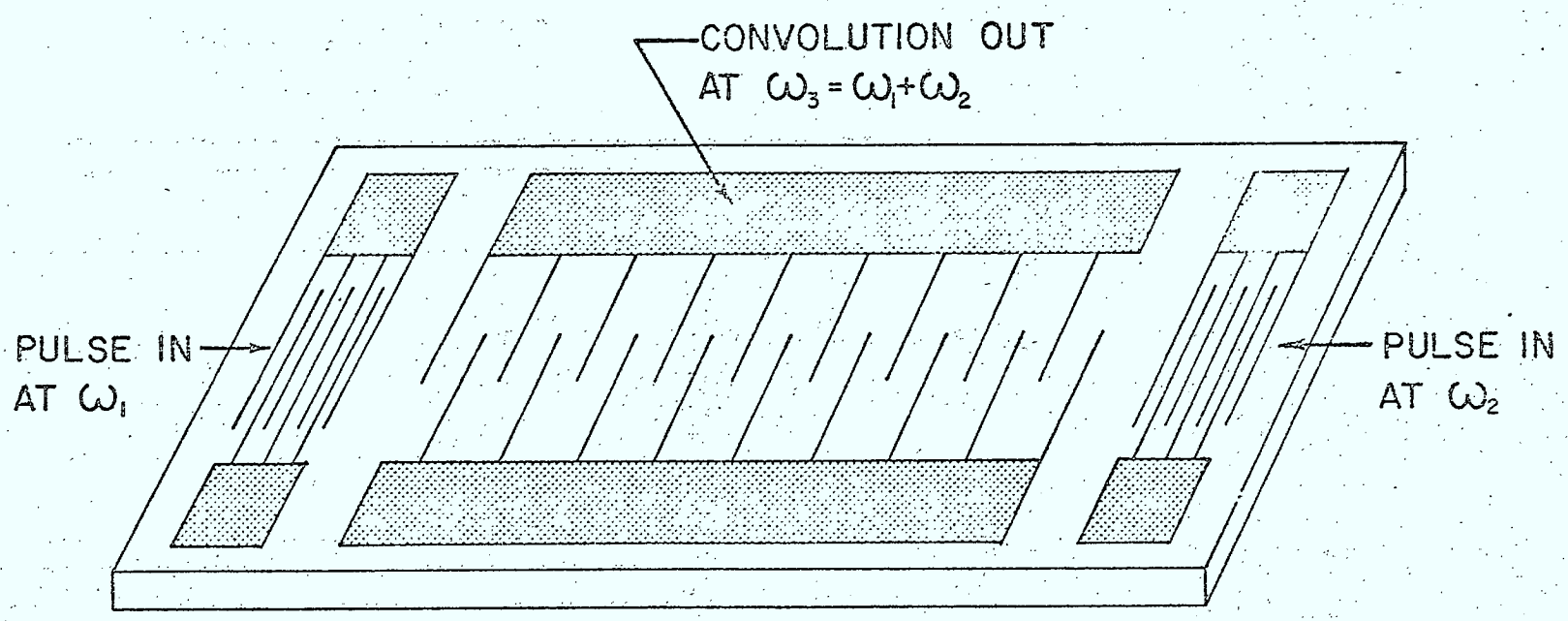


Fig. 13.2 Schematic of surface-wave convolver with interdigital integrating electrode.

If each signal has a duration less than the transit time along the length of the output transducer then the integration limits can be taken as $\pm\infty$ and using the transformation $\tau = t - x/v$ the output becomes

$$C(t) = -vB \int_{-\infty}^{\infty} F(\tau)\, G(2t - \tau)\, d\tau \qquad (13.1)$$

which is Eq. 13.1 repeated. As mentioned earlier this differs from the mathematical definition of convolution by a time compression factor of two due to the fact that the relative velocity of the signals is twice the acoustic velocity v. The convolution takes place in real-time and the F and G functions can be chosen quite arbitrarily as long as their duration is less than the transit time under the integrating transducer. Analog signals as well as digital codes can thus be processed although the device is really an analog device. Note also that the output transducer is broad band since any simultaneous change $\Delta\omega$ in both input frequencies keeps k_3 unchanged.

An important special case occurs when $\omega_1 = \omega_2 = \omega$ and the output frequency is 2ω with $k_3 = 0$, which implies that the sum frequency signal has no spatial variation and can be detected by plate electrodes on the top and bottom surfaces of the delay line. This structure was discussed earlier and is shown in Fig. 13.1.

A circuit model for this type of convolver is obtained by means of the reciprocity theorem to give the device open circuit voltage and Thévenin impedance. The maximum open circuit voltage can be written as (42)

$$V_{oc} = \frac{M}{W_3} (P_{ac1} P_{ac2})^{1/2} , \qquad (13.3)$$

where P_{ac1} and P_{ac2} are the acoustic input powers, W_3 is the width of the output transducer and M is the non-linear proportionality parameter and is clearly a figure of merit for any convolver, see (72) for a table of M for common materials. Note that this is independent of the output transducer length and the thickness of the piezoelectric crystal. The Thévenin impedance is essentially the capacitance to ground of the centre pick up transducer.

It is convenient to define a parameter which gives a measure of the device efficiently. The overall performance is measured in terms of a bilinear coefficient F_T defined by

$$F_T = P_3 / (P_1 P_2) \tag{13.4}$$

with P_3 the power delivered to a 50 Ω load, P_1 and P_2 the electric powers supplied to ports 1 and 2 . Another suitable measure of nonlinear coupling is the internal bilinear coefficient F_{INT} defined in terms of internal acoustic powers

$$F_{INT} = P_3 / (P_{ac1} P_{ac2}) , \tag{13.5}$$

this internal coefficient is more suitable if we wish to compare the strength of nonlinear coupling in different devices without taking into account losses and transducer efficiency, both F_T and F_{INT} are often expressed in dB . In systems applications not only will these bilinearity factors be important but so will the bandwidth, maximum delay or integration time, and the time-bandwidth product.

The convolution devices depicted in Figs.13.1 or 13.2 can also be used to carry out a number of other signal processing functions such as variable time delay,

correlation, time inversion, and time expansion and compression by a factor of 2.

Variable time delay is obtained by applying a delta function at time t_1 to port 1 so that in Eq.13.1 $F(\tau) = \delta(\tau - t_1)$ in which case $C(t) = -v B G(2t - t_1)$ so that a delayed replica of the signal $G(t)$ applied to port 2 appears at port 3 which is time compressed by a factor of two.

Correlation results when signals F and G are applied to ports 1 and 3 respectively. For this combination of signals it can be shown that the product produces a signal at ω_2 which propagates back towards port 1 and which can thus be detected by a transducer tuned to ω_2 placed at the same end as the input transducer. For the degenerate convolver of Fig. 13.1 since $\omega_2 = \omega_1$ the same transducer launches the signal F and detects the modulation signal given by Eq. 13.2, which as mentioned in Section 13.1 is the correlation of F and G time expanded by two.

Time reversal is obtained by applying a delta function at port 3 so that in Eq. 13.2 $G(\tau) = \delta(\tau)$ in which case $C(t) = -\mathfrak{L} v F(-t)$ and a time reversed replica of the signal $F(t)$ applied to port 1 appears at port 1.

All the functions which the surface-wave convolver can perform are summarized in Table 13.1.

13.4 Performance of Convolution Devices

There are three types of surface-wave convolvers which have been investigated and for which performance data are available. The first type called the non-

TABLE 13.1

BASIC FUNCTIONS PROCESSED BY A PARAMETRIC SURFACE-WAVE SIGNAL PROCESSOR

Function	Port 1	Port 2	Port 3	Output
Convolution	Input $F(t)$ Traveling Wave	Input $G(t)$ Traveling Wave	—	Port 3 $\int_{-\infty}^{\infty} F(\tau)\, G(2t-\tau)\, d\tau$ Standing Wave
Electronic Time Delay and Time Compression	Input $\delta(t-t_1)$ Traveling Wave	Input $G(t)$ Traveling Wave	—	Port 3 $G(2t-t_1)$ Standing Wave
Correlation	Input $F(t)$ Traveling Wave	—	Input $G(t)$ Standing Wave	Port 1 $\int_{-\infty}^{\infty} F(\tau)\, G(2\tau-t)\, d\tau$ Traveling Wave
Time Inversion	Input $F(t)$ Traveling Wave	—	Input $\delta(t)$ Standing Wave	Port 1 $F(-t)$ Traveling Wave
Time Stretching	Input $\delta(t)$ Traveling Wave	—	Input $G(t)$ Standing Wave	Port 1 $G(\frac{1}{2}t)$ Traveling Wave

linear acoustic convolver utilizes the material nonlinearity of the piezoelectric delay line and it is this type which was used as the basis for the discussions in Sections 13.1 and 13.3. This type has tended to be uncomfortably inefficient and recently much more efficient convolvers have been fabricated that use the nonlinearities of a semiconductor closely coupled to the RF fields of a piezoelectric delay line. The third type of nonlinear processor utilizes an external diode array connected to a series of equispaced interdigital taps. The mixing of signals received by the diode-taps is summed to create a sampled data type of convolution.

Nonlinear Acoustic Convolver

Since this convolver has already been discussed in some detail in Section 13.3 it suffices to just summarize typical performances obtained with such devices using YZ Li Nb 0_3.

Interaction time delays (Convolution interval)	~	10 μ S
Bandwidths	~	100 MHz
Time bandwidth product	~	1000
Bilinearity Coefficient F_T	~	- 80 dB
Figure of Merit M	~	1.2×10^{-4} v m / w
Dynamic range	~	60 dB

Note that although the insertion loss of these convolvers is very high the time bandwidth product is rather impressive.

Semiconductor Coupled Convolvers

There are two ways which have been used to couple the piezoelectric fields accompanying a surface-wave to the carriers in a semiconductor in order to obtain a strong nonlinearity. The simplest method is to bring a semiconductor wafer in close proximity to the delay line much as in the separated medium amplifier discussed in Chapter XII. This requires some method for maintaining a uniform gap between semiconductor and surface-wave delay line, such a convolver in which the air gap is fixed by a space rail is shown in Fig. 13.3. A novel approach has been devised (G6, p.142) for air gap separation which consists in supporting the silicon with a pseudo-random distribution of posts having 3 μm diameters separated by about 200 μm thus minimizing both scattering of the wave at the operating wavelength of 17 μm, and any flexing of the 0.2 mm thick silicon strip. The posts are made by rendering the pattern in photoresist on $Li\ Nb\ O_3$ by a standard photolithographic technique and then ion-beam etching the exposed surface of the crystal by a few thousand angstrom and removing the photoresist. By means of a rubber pad, uniform pressure can be applied resulting in a structure which is mechanically stable and rugged. A schematic of such a convolver is shown in Fig. 13.4 giving a sectional and a bottom view of the device.

The basic interaction process which gives rise to the nonlinearity is due to the normal electric field E_n from the acoustic wave giving rise to a depletion layer (41) at the semiconductor surface. A voltage is generated across this depletion layer which is proportional to E_n^2. In the presence of two signals each at a frequency

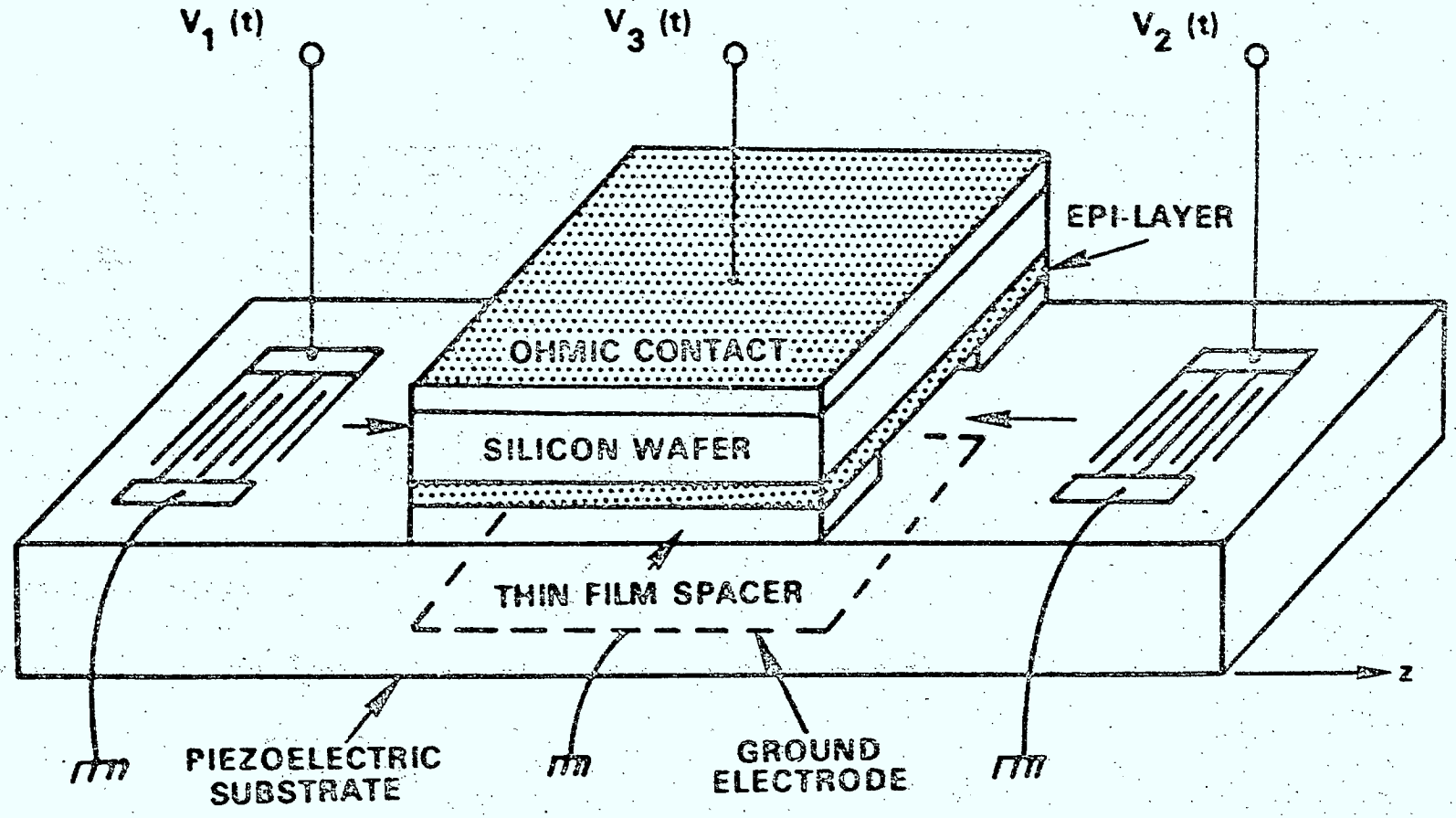


Fig. 13.3 Diagram of separated-medium semiconductor-coupled convolver.

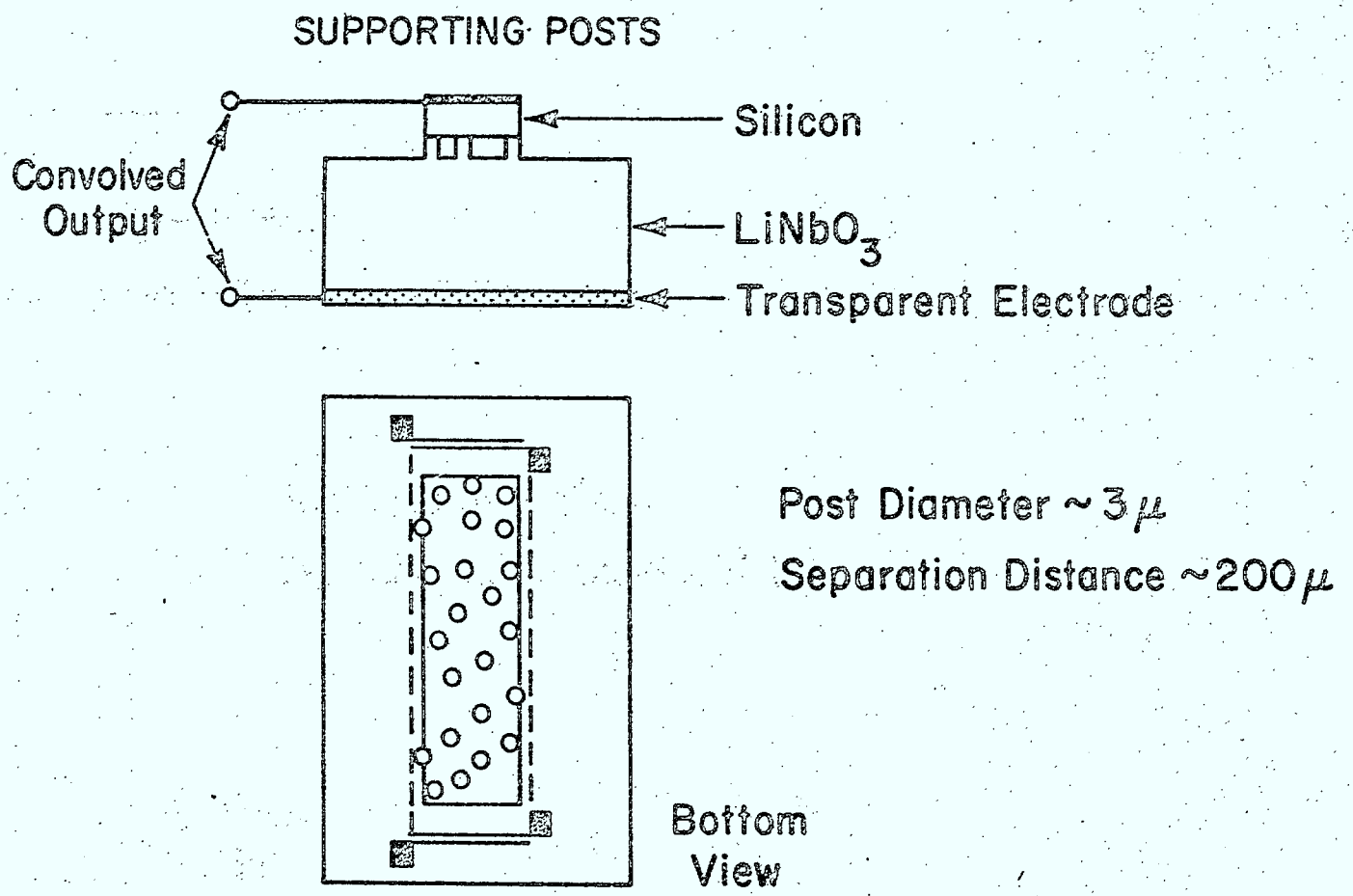


Fig. 13.4 Schematic of post-supported Si – Li Nb 0_3 convolver (G6, p.142).

ω producing fields E_1 and E_2 normal to the semiconductor surface, there will be a second harmonic potential proportional to the product $E_1 E_2$ and a DC potential proportional to $(E_1^2 + E_2^2)$ developed across the depletion layer. This DC potential is called the "transverse acoustoelectric effect" and will only be important at high signal levels.

The detailed theory for this air-gap convolver is described in (41) and the best results obtained with this device as reported in (G6, p.142) are given below for a 10 Ω. cm resistivity silicon device 1 cm long and .75 mm wide.

Convolution interval	9 μ S
Bandwidth	65 MHz
Time Bandwidth product	> 500
Centre frequency	200 MHz
Dynamic Range	55 dB
Bilinearity Coefficient F_T	- 60 dB
Figure of Merit M	Not available

The dynamic range given is with input power levels of 20 dBm at which levels no saturation effects are observed. The design techniques and fabrication methods can be extended to yield convolvers with time bandwidths of 1000 and bandwidths in excess of 100 MHz, and by making the silicon strip width narrower the bilinear coefficient F_T can be increased.

Strip-Coupled Convolvers

Recently a second type of semiconductor coupled convolver, the strip-coupled convolvers, which uses a multistrip coupler to couple the acoustic surface-wave to a semiconductor located outside the acoustic beam path has been described (G6,p.145). This device has a figure of merit which is 20 dB better than the best air-gap convolvers. The configuration employed is shown in Fig. 13.5. In this convolver a large number of strips are deposited on the Li Nb 0_3 spanning and extending beyond the acoustic beam path. An n on n^+ silicon wafer located outside the beam path is pushed down on the metal strips, the side in contact with the metal strips has an oxided layer of thickness h and the overlap width is a . In this way the region of overlap of each strip with the silicon forms a small MOS varactor diode whose area is controlled by varying the overlap dimension a , the individual diodes behave as nonlinear capacitors and hence as mixers for two input signals. When two input signals of frequency ω are applied an output signal at 2ω will be generated between the semiconductor and a metal electrode on the opposite surface of the Li Nb 0_3 just as in the degenerate convolver described earlier.

The interaction process which gives rise to the nonlinearity is identical to the gap coupled convolver. The normal component of the RF fields of the acoustic wave giving rise to a depletion layer (41) across which a voltage proportional to the square of the field is generated. When two signals each at a frequency ω are present a second harmonic voltage proportional to the product is developed across the depletion layer.

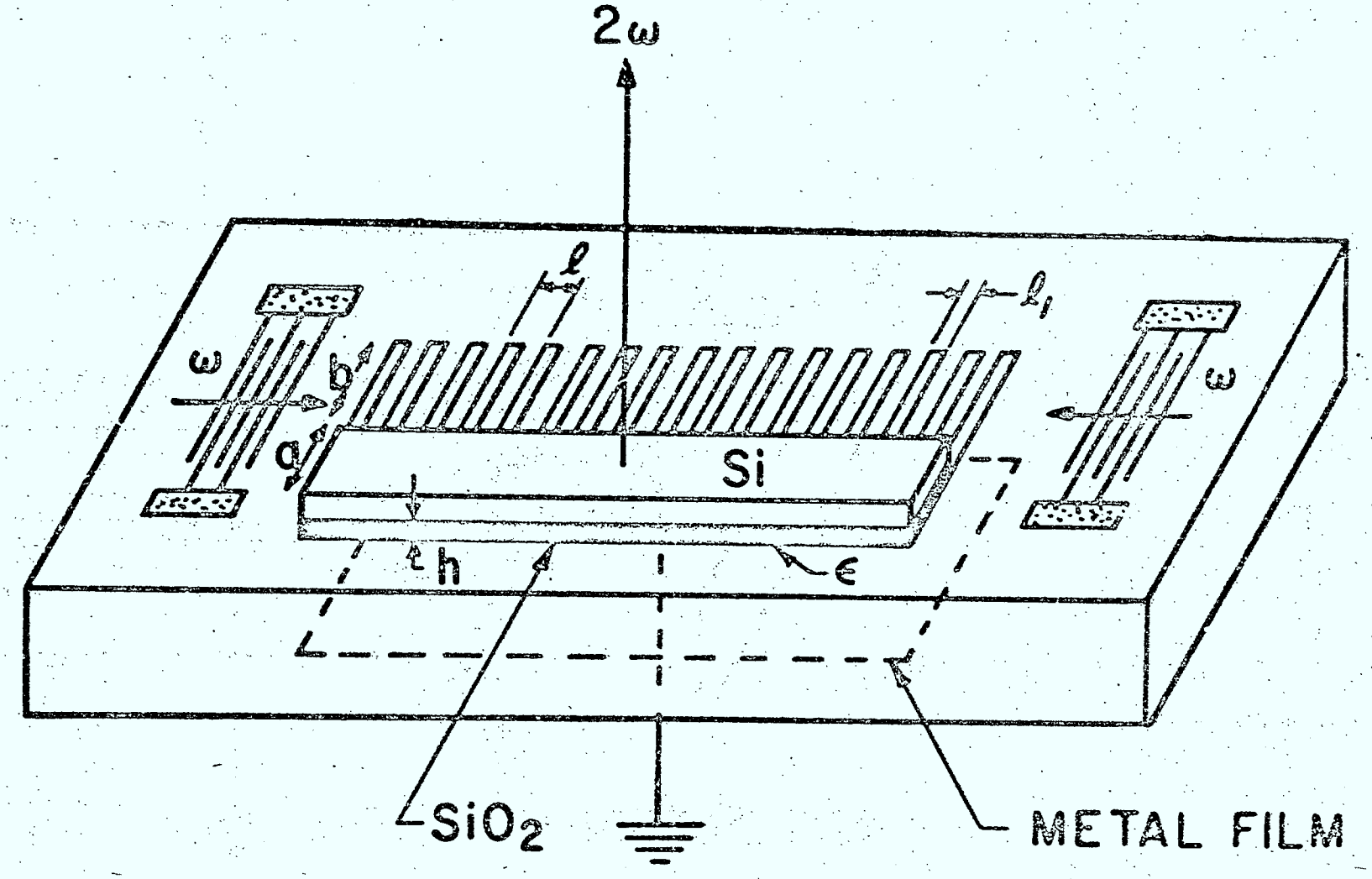


Fig. 13.5 Schematic of strip-coupled convolver (G6, p.145) .

The detailed theory for this strip coupled convolver is described in (G6,p.145). The most efficient device was obtained with a Li Nb 0_3 substrate and 7 Ω cm n on n^+ epitaxial silicon with an effective oxide layer of 1300 Å. The strip coupler used was designed with three strips per wavelength at 42.5 MHz, a metallization ratio $l_1 / l = 0.8$ (see Fig. 13.5), and the ratio $a / b = 0.3$. For small signal levels (below 0 dBm) the bilinear coefficient for this convolver which represents the external device efficiency is $F_T = -23$ dB. The nonlinearity strength is measured by the figure of merit M, as defined by Eq. 13.3. This factor represents internal device efficiency and is independent of transducer losses, output tuning and acoustic beam width. The value $M = 0.6$ Vm / w for this device is one hundred times larger than the best air-gap convolver (72). One disadvantage of this device resulting in reduced dynamic range is that the bilinearity coefficient is power dependent since, as a result of the DC potential generated acoustoelectrically, the carriers are depleted resulting in an effective oxide thickness which is greater than h. When this acoustoelectric effect is included in the analysis the results predicted theoretically are in excellent agreement with experimental data both at low signal levels and at large signal levels. For the same device described earlier at equal input power levels of 20 dBm $F_T = -39$ dBm, and $M = .09$ Vm / w.

To summarize, the strip coupled convolver has a figure of merit which is almost 20 dB better than any other Si - Li Nb 0_3 convolver constructed. It is expected that the semiconductor can probably be bonded to the metal strips hence making a monolithic device.

Diode Convolvers

The third type of convolver utilizes a tapped delay line consisting of a series of equispaced interdigital taps connected to an external array of semiconductor diodes. It will be recognized that this device is similar in principle to the strip coupled convolver described earlier, the main difference is that the use of external diodes requires a large number of connecting leads and as a result the tap spacing cannot be made too small and the device bandwidth is limited. For devices under current construction the usable bandwidth is no better than 30 MHz (G4,p.73) and unless a combined piezoelectric / semiconductor thin film technology such as Al N and Si on sapphire is developed which will allow the diodes and delay lines to be fabricated on a common substrate it is unlikely that the diode convolver can compete with any of the other semiconductor convolvers which have been described.

In summary this chapter has outlined the operating principles and operating characteristics for three types of nonlinear surface-wave devices: the nonlinear acoustic convolver, the semiconductor air gap and strip coupled convolver, and the diode convolver. The semiconductor convolvers are the most promising with a time bandwidth product potential in excess of 1000, and an accuracy and dynamic range consistent with many systems applications.

XIV. CONCLUSIONS

This report has reviewed the fundamentals of surface-wave devices and it is evident that such devices now constitute an established technology to provide a viable alternative to other forms of hardware for a variety of signal-processing roles in the frequency range 1 MHz to 2 GHz. The upper frequency limit is set by losses and the present fabrication technology, the lower limit is set by available size of materials and bandwidth considerations. The bandwidths and operating frequencies which are currently available are appropriate to IF signal processing requirements of many radar and communication systems. The surface-wave components are extremely small and rugged, and the fabrication techniques are compatible with established planar technology; usually consisting of no more than a metallization pattern on a polished single-crystal substrate. The initial impact of surface-wave devices has been in high-performance radar where they have made important contributions in matched filtering and pulse compression due to their competitive performance, reliability, reproducibility and cost effectiveness.

There are innumerable demonstrated and proposed applications of surface-wave devices, almost all of which have a strong competitor in another technology such as transistor digital logic or the newer charge-coupled devices. Thus the expansion of surface-wave devices into the potentially available fields depends on factors other than fundamental ability to carry out some system function. Such deciding factors (cost, reliability, size, reproducibility, availability, second-order effects) are critically dependent on the amount of development effort expended on

this technology relative to the competitors. It is not evident at this time how much of a gamble is involved in large-scale commitment to surface-wave development, but there is no doubt that surface-wave devices will play an important role in many communication systems of the coming generation.

The complementary nature of surface-wave and microelectronic integrated circuits from the stand points of electronic signal processing, fabrication, and computer-aided design should be regarded as an important attribute for surface-wave devices. This complementarity is particularly evident in the hybrid integration of surface-wave delay lines with silicon IC amplifiers and diodes to realize programmable filters and high frequency oscillators.

The main inherent properties that will no doubt guarantee a continuing role for SAW devices are:

- frequency selective operation covering the HF, VHF, UHF and low microwave frequencies,
- controlled sampling of real-time signals,
- high dynamic range (> 50 dB) passive signal processing,
- simple manufacturing technique, microelectronic compatibility, reliability,
- demanding specifications can be met using computer-aided design since the underlying physics is well understood.

Space has of necessity precluded discussion of many systems applications and all the signal processing versatility of SAW devices. In digital signal processing some examples are recirculating memories for computers and ECM applications, data and bandwidth compression schemes, discrete z-transform and discrete Fourier Transform realization, and picture processing applications. Another important developing area is in electronic scanning of optical and acoustical images. Table A 3.1 in Appendix III prepared by J. Collins (presented at the 23rd MRI Symposium, New York, 1974) shows one form of listing of surface-wave prototype devices, their derived developments and potential areas of applications. The status of surface-wave components in the present radar system context is summarized in the interim conclusions of the February 1974 report of Group KAG3 of TTCP entitled "The Applications of SAW devices to Radar Signal Processing," which we include as Appendix III.

Many new ideas associated with surface-waves are being pursued and there is ample room for further research, innovation, and new devices. Surface-wave convolution devices are worth further investigation as they seem to offer the most promise as electronically programmable matched filters. Methods are being studied for storing a finite duration signal on the surface of a piezoelectric crystal for several minutes in an attempt to fabricate waveform correlators. The processing of information in 2 dimensions is also being studied, the immediate applications being in radar doppler processors and in spectral analysis.

The guide to the bibliography, Appendix I, has been written for the reader who wishes to study some of the aspects of this new technology without getting overwhelmed by huge lists of reports and papers. The guide selects for the area as a whole a set of survey papers, and for each topic of the report a restricted set of papers which form the basis of any study of that topic.

Finally we should mention the Canadian Workshop on Acoustic Surface-Waves which was held in Ottawa under the sponsorship of Communications Research Centre. This meeting revealed that a number of University, Industrial, and Government Organizations in Canada are aware of the new technology and its potential and that much expertise already exists in Canada. For details concerning this Workshop consult Communications Research Centre Report No. 1260, (1974).

APPENDIX I

GUIDE TO BIBLIOGRAPHY

It has been noted several times that surface waves form the basis of a reasonably mature technology based on a thorough understanding of the underlying physics. To reach this state over the half dozen or so years involved has required, or at least produced, an exponential growth in the published literature. It is thus difficult for the neophyte to find his way through a long list of frequently contradicting reports and papers in a field where no directly applicable textbook exists. We attempt here to give a little guidance to the newcomer to the field in attacking this literature.

The entries G1 to G11 in the bibliography are review works of one form or another. Our hypothetical newcomer would do well to start with the summary by Kino and Matthews(G9) and follow this with the later and more detailed review by Maines and Paige (G10), the two introduce much of the material covered in this report and many of the device principles are outlined. An earlier review by White (G11) covers much of the background material on surface-wave principles but it was published before the systems oriented devices became common, however its extensive bibliography also remains very useful.

The details of surface-wave propagation on crystalline substrates are covered at length in Auld (G1), Farnell (G2) and Farnell and Adler (G3).

A reasonable feeling for the state of the art in surface-wave devices in 1969 can be obtained from the special issue of the IEEE Proceedings on Microwave Theory and Techniques (G7); and a later special issue on signal processing with surface-waves published jointly by the IEEE Groups on Microwave Theory and Techniques and on Sonics and Ultrasonics (G8) illustrates many of the applications of surface-wave devices.

The major annual conference in the surface-wave field is the IEEE Symposium on Sonics and Ultrasonics for which the proceedings were published in 1972 (G5) and 1973 (G6) and provide not only a reference for the current state of the art but also of the international interest in this field. There was also an important ad hoc conference in 1973 at Aviemore in Scotland (G4) which attempted with a great deal of success to present the viewpoints of both the systems engineers and the surface-wave device designers.

Surface-wave transducers are discussed in considerable depth in three survey papers by Smith et al (85, 86, 87) which contain the basic analysis and design concepts based on the equivalent circuit model. Hartmann et al(27) show that the circuit model approach can be used to predict the performance of unidirectional transducers. For a theoretical analysis of transducers based on solutions of the wave equation the reader should consult Tseng (99), Engan (19), and Emtage(18). The effect of finger resistance has been analyzed recently by Lakin (46). Experimental data and performance specifications can be found in the Ultrasonics Symposium Proceedings (G5, G6).

The analysis and design of band-pass filters are reviewed by Tancrell (95, 96) and Hartmann et al(26), and the transversal filter viewpoint is treated by Matthai (63). Basic filter design philosophy is discussed by Hurlburt (32) and the use of surface-wave filters in spectral analysis has been studied by Guyot (22). Computer analysis programs have been developed by Tancrell and Sandy (97), and Slobodnik (81); and McClellan et al's (64) digital filter design techniques are directly applicable to surface-wave filters. Many filters have been fabricated and their measured performance compared with theoretical calculations by De Vries et al for T.V. filters (16), Cheek et al for low shape-factor filters (11), Gerard et al for phase correction of dispersive filters using dummy electrodes (21), and Smith et al for chirp filters (87, 88).

The details of phase coded devices are covered by Bell et al (5), Hunsinger (30), and Jones et al (35). Programmable devices are reviewed by Staples and Claiborne (91). Experimental data on these devices is given by Hunsinger and Franck (31), De Vito et al (15), O'Clock et al (69), and Claiborne et al (12). Performance limitations due to reflections from metal electrodes are discussed by Jones et al (36) whereas the effect of temperature and doppler shift has been studied in detail by Carr et al (10). Considerable experimental data will be found in the two reports by Judd et al (37).

Reflective array compressors are described by Williamson and Smith (106), oscillators by Davies and Lawrence (14), and multistrip couplers by Marshall et al (59, 60) and Maines and Paige (G10).

Most of the important materials and their properties are listed in Slobodnik (79), Szabo and Slobodnik (94, 83), and Schultz and Matsinger (75). Small temperature coefficient materials are discussed by Schultz and Holland (74) and piezoelectric films are evaluated by Hickernell for ZnO (28) and by Duffy et al for Al N (17).

An up-to-date survey by Bahr(G4, p.22) describes most of the current fabrication methods. Smith et al (84) and Angerstein and Voges (3) give details of photoresist methods of pattern definition. Broers and Hatzakis (8) and Ozdemir et al (73) discuss methods which use electron beam and scanning electron microscopes, while Spears and Smith outline the x-ray lithographic process (89).

The details of diffraction associated with surface-wave devices are covered in Kharusi and Farnell (38), Szabo and Slobodnik (93), Mason and Ash(62), and Mason (61).

Acoustic waveguides are surveyed by Stern (92), and the theory for surface guides covered in Lagasse et al (45), Tiersten (98), and Tu and Farnell (100). Hughes (29) discusses the $\Delta v / v$ type of guiding and Oliner et al (70) have applied a network formalism to waveguiding problems.

Surface-wave reflections due to assorted types of discontinuities are covered by Munasinghe and Farnell (66), Li and Melngailis (G6,p.503), the multistrip type of reflector by Marshall et al (60).

Surface-wave amplifier theory and experiments are reviewed in a paper by Lakin and Shaw (48), and the detailed theoretical analysis is given in Kino and Reeder (43), Ingebrigtsen (33), and Lakin (47). The theory of multistrip coupled amplifiers is covered in Kino (39) and Maerfeld and Tournois (G5,p.171). Experimental results for various types of amplifiers have been reported by Ludvik and Quate (54), Kino and Coldren (40), Coldren and Kino (13), and Hanebrekke and Ingebrigtsen (25).

Nonlinear effects observed in surface-wave devices are reported by Lopen (53), Slobodnik (80), and Lean and his co-workers (49, 50). Theoretical analysis of harmonic generation have been developed by Tiersten (G6,p.244) and Vella and Stegeman (101).

Surface-wave convolvers are discussed in a review paper by Kino et al (42). The theory for such devices is covered in Luukkala and Surakka (55) and in Kino and Gauthier (41). Data on materials for convolvers is given in Otto and Moll (71), and the most recent performance characteristics of these devices is reported by Smith et al (G6,p.142) for a semiconductor coupled convolver, and by Shreve and Kino for a strip-coupled convolver (G6,p.145).

APPENDIX II

FOUR REPRESENTATIVE PACKAGED DEVICES

A 2.1 - Surface-Wave Bandpass Filter Module (11)

Substrate	- 20° Rotated Y - cut X - Prop Crystalline Quartz
Center Frequency	168 Mhz
Bandwidth	12 Mhz
Midband Insertion Loss	14 dB (50 Ω system)
Main Lobe / First Side Lobe	12.6 dB
Skirt Slope	80 dB / Mhz
Shape Factor	1.32 (3 dB B.W. / 40 dB B.W.)
Bandpass Ripples	~ 3 dB
Phase Response	< 10° deviation from Linear Phase
Temp. Sensitivity of Center Frequency Drift	< 200 KHz from - 55° to 125°C
Net Gain using Integrated Amplifier Module	6 dB

A 2.2 - Chirp Filter (88)

Substrate	Y - cut, Z - Prop. Li Nb 0_3
Center Frequency	300 Mhz
Bandwidth	100 Mhz
Midband Insertion Loss	30 dB
Compression Ratio	1000 : 1
Pass Band	~ Ideal theoretical shape
Sidelobe Suppression	13 dB unweighted 27 dB Hamming weighted
Bandpass Ripples	± 0.8 dB
Phase Response	< 20° deviation from ideal quadratic function
Triple Transit Echo Suppression	> 50 dB
Null Width	20 nsec unweighted 38 nsec Hamming weighted

A 2.3 - 127 Chip Phase Coded Delay Line (37)

	Current Performance	Projected Performance
Substrate	ST - cut X - Prop. - Quartz	
Center Frequency	120 MHz	150 MHz
Bandwidth (Chip Rate)	5 MHz	10 MHz
Midband Insertion Loss	36 dB	-
Recompressed Pulse/Sidelobe	22 dB	Within 1 dB of Theoretical for Selected Code
Pulse Length	25.4 μ sec	12.7 μ sec
Compressed Pulse	0.4 μ sec	0.2 μ sec

A 2.4 - Programmable 127 Chip Phase Coded Delay Line (23, 24)

	Current Performance	Projected Performance
Substrate	ST - cut, X - Prop. Quartz and Silicon- on- Sapphire Control Circuits	
Center Frequency	120 MHz	435 MHz
Bandwidth (Chip Rate)	10 MHz	20 MHz
Midband Insertion Loss	27 dB	-
Correlation Peak to Sidelobe Ratio	20 dB	Within 1 dB of Theoretical for Selected Code over Temperature Range
Correlation Null Width	200 ns	100 ns
Tap Spacing	100 ns	50 ns
Operating Temperature Range	- 25° to + 85°C	0 to 50° C

APPENDIX III

Application of Surface Acoustic Wave (SAW) Devices to Radar Signal Processing*

"It was found that the understanding of SAW phenomena is at the stage where some new systems are being designed around acoustic signal processing devices whose performance can be anticipated with confidence; other systems are being experimentally retro-fitted with SAW components with the expectation of improving their performance, reducing size etc. while there is a third, more speculative, generation of devices which may provide the means of performing complex processing functions (correlation, Fourier transformation) on a very general range of signals. . . .

The chief impact of SAW devices on radar signal processing has been to put into the hands of the system engineer a means of realising very precisely and economically some of the long understood principles of optimum modulation and reception filtering. The passive nature of the SAW processer, its compactness, robustness and reproducibility are secondary but of course very important attributes. For a fixed processing function, the advance of digital and other techniques, employing for instance charge transfer devices or magnetostatic waves are unlikely to compete for some time with the precision and convenience of using the simple structure of a planar SAW component. SAW device construction is often trivial by modern standards of electronic circuit fabrication since it is merely a reproducible metallised pattern on a flat substrate. Its behaviour is well understood and conforms very well to design parameters which are usually introduced as data for a computer aided design

---

* Interim conclusions from report of this title by Group KAG3 of the Technical Cooperation Program (TTCP), February 1974.

procedure. The range of bandwidth and delay times achievable with good dynamic range is very significant for radar purposes: indeed the spatial resolution of many radars is likely to improve through the ease of handling large bandwidths with good phase control by SAW components. Thus, for the generation and optimum filtering of fixed waveforms with TB products below 1000, T below 100 μs and B in the range of 10 KHz to 500 MHz, SAW devices are likely to be pre-eminent for the foreseeable future. Electron beam fabrication is extending the bounds on time and bandwidth beyond these limits; for example 1 ms delays and 1 GHz bandwidths can already be anticipated. Since SAW filters are small system components, it would be simple to introduce modulation agility by switching between alternative devices in the IF transmission / reception circuits.

Where greater flexibility and longer delays are required, SAW meets stiffer competition from time-quantized systems. However it is possible for SAW devices to utilise digital pseudo-random coding and decoding techniques and these could prove competitive for secondary radar and missile fuzing systems operating at high bit rates and where interference / ECM resistance, size, power consumption and robustness are crucial. Polyphase coding flexibility is also possible for SAW devices provided with programmable tapping points. Similarly switched devices are used to generate the transmitter coding and to provide matched filter reception. The device complexity is roughly comparable with a one-bit digital pulse compression system but in contrast, because it is a linear processer, it is not confused by the almost inevitably overlapping codes in the received signal due to clutter, jamming, multiple targets or multipath.

The implications of SAW for ECCM are likely to be important since complex wideband modulations can be employed even in small radars to combat interference and jamming. However a note of caution is needed since, for wide spectrum techniques to succeed against deliberate jamming and possibly against interference from similar friendly radars, a bandwidth of > 1 GHz may be needed and this is not presently available from SAW devices. This contrasts with communications systems where only one-way transmissions are involved. In radar the two-way (σ / R^4) echo has to compete with one-way $(1 / R^2)$ interference.

It seems appropriate to draw attention to the possibilities of employing SAW in an ECM role. We have already mentioned variable delay lines as being well suited for (expendable) deception devices using range gate stealing. These might require the aid of a compressive, or microscan, intercept receiver to determine the RF being used, for which dispersive SAW lines or contiguous filter banks offer a powerful technique.

The SAW contribution to IF circuit techniques is likely to be significant in radar (as elsewhere) since simple bandpass filters with well defined amplitude and phase characteristics can be designed and reproduced accurately. Also stable, but readily FM'd high frequency SAW oscillators which are small and rugged should find several areas of application.

SAW convolvers and correlators, which depend on non-linear effects, are still in the research phase. In particular to achieve correlation using a convolver requires a time inversion for which there is no well established technique and

asynchronous operation is still difficult. Also, it is not readily apparent that the kind of flexibility they offer is useful in a radar context: most radar designers are likely to be content with a prescribed form of modulation (even when they require it to have random properties) and do not need complete generality in the choice of waveform. Unlike digital or charge transfer devices, a SAW correlator could not be used in a range of systems of widely differing bandwidth and storage time parameters because T and B are not flexible.

There are radar signal processing problems for which SAW techniques do not offer the best approach. This is chiefly evident where integration times measured in msec are involved. Thus MTI and doppler processing, which typically require a frequency resolution of 0.1 to 1 KHz and therefore a storage time measured in msec, are at present best achieved using sampled data techniques while research proceeds on long SAW delay lines.

A long term future seems assured for SAW in radar waveform generation and filtering. A range of significant radar applications where SAW is less likely to be dominant, for instance oscillators, is also foreseeable, together with some special requirements which may admit of a SAW solution, as for example generating a synthetic clutter signal to ensure reliable detection by a clutter locked MTI system.

As SAW devices improve, more will be demanded of the associated electronics in order not to limit system performance by non-linearities (in amplitude and phase) of the amplification and mixing stages. This factor must not be overlooked in the quest for lower sidelobes and increased dynamic range in SAW devices.

Its importance is particularly evident in the design of receivers for pulse compression radar. The achieved range resolution can fall short of the equivalent short pulse radar for multiple targets and strong clutter due to non-linearity before compression stage. This leads to very stringent requirements on the linearity and dynamic range of the receiver front end to realise the full capability of the SAW pulse compressor.

It should be emphasized that the investment in SAW technology, compared to IC technology, has been very low while the return on that investment has been very high. For military radar systems in particular, it offers an alternative way to sophisticated signal processing with a minimum of digital active integrated circuits. This alternative is potentially cheaper, more reliable, and less power-consuming. The current level of government research support in the SAW field is estimated to be less than $2.5 million per year (in the US) and this effort has application in EW, communications, navigation and other fields, besides radar. The further development of SAW devices are mainly for applications in the military field so that, with a few exceptions, there is little commercial incentive to advance the technology quickly. This situation justifies continued government support for research into SAW materials, materials processing, device concepts and device design in order to achieve greater economy in the design of radar systems.

Finally, it is clear that SAW devices have made a significant contribution to radar signal processing, a tendency which is certain to be accentuated in the next few years. The research effort is well in touch with the applications needs, the major outstanding uncertainty being about the best ways that SAW devices can complement other emerging technologies whose limitations and strengths are not so well established . . . . "

TABLE A 3.1

SURFACE WAVE DEVICES AND POTENTIAL AREAS OF APPLICATION

SURFACE WAVE DEVICES		POTENTIAL AREAS OF APPLICATION
BASIC	DERIVATIVES	
INTERDIGITAL TRANSDUCER	Frequency Filter	Color TV, Integrated IF's; Satellite Multiplexers; ECM
	Inverse Filter	Clutter Dominated Radar
	Discriminator	Analog FM Demodulator
	Oscillator	Stable Sources VHF to Microwave; Specialised Comms; Frequency Synthesiser
DELAY LINE		Altimetry, Analog MTI Radar, Comms Path Length Equaliser; Time Ordering
	Tapped	Radar Clutter Reference; ECM; SSR
	Bit Matched Filter	Digital Comms PSK Demodulator
	Acoustic Amplifier	Long Delay for TV Frame Storage
	PSK Matched Filter	Spread Spectrum Comms; IFF; Integrated Comms; Navigation and Identification
	PN PSK Generator	Link Analysis; Spread Spectrum Comms
	Digital	Recirculating Memory; Data and Bandwidth Compression; Sonar
	Touch Sensitive	Man-Machine Interface; Glass Intrusion
MULTI-STRIP COUPLER	Improved Devices	Signal Routing
DISPERSER		High Resolution Radar; Navigation; Group Delay Equalisers
	Compressive Receiver	Spectral Analysis
	Variable Delay Line	Range Calibration; Target Simulation; Electronic Timing; ECM
NON-LINEAR	Convolver	Spread Spectrum Comms; Radio Astronomy; FFT Processor
	Acoustic Imaging	Non-Destructive Testing
	Acousto-Optic	Displays, Cameras
WAVEGUIDE		Signal Routing
	Acoustic IC's	Subminiature Signal Processing Sub-Systems

APPENDIX IV

FIGURE CAPTIONS

Page		
6	Fig. 2.1	Prototype geometry for surface-wave device.
9	Fig. 2.2	Surface-wave velocity on isotropic substrates. v_t = transverse or shear bulk-wave velocity. v_l = longitudinal bulk-wave velocity.
10	Fig. 2.3	Amplitude of displacement components as a function of depth. Displacement components are in phase quadrature.
11	Fig. 2.4	Distortion of a square grid in sagittal plane due to surface-wave. Displacement amplitudes are exaggerated.
12	Fig. 2.5	Mechanical Poynting vector, power flow per unit area perpendicular to sagittal plane, as a function of depth below surface.
16	Fig. 3.1	(a) Geometry of an interdigital surface-wave transducer. (b) Sketch of electric field in the substrate.
17	Fig. 3.2	Electric field components (arbitrary units) at the surface and $L/4$ below the surface for one section of an interdigital transducer.
19	Fig. 3.3	(a) Cross-section of an interdigital transducer. (b) Side-electroded bar model of one section of transducer. (c) Mason equivalent circuit for length L of transducer based on (b). Z_o is mechanical impedance of bar, θ is transit angle $= \pi \omega / \omega_o = 2\pi L / \lambda$, ω_o is resonant frequency $= \pi v / L$, and C is static capacity of section.

Page

21 Fig. 3.4 Electrical equivalent circuit for ith section of transducer. R_o defined in Eq. 3.1, θ in Fig. 3.3.

24 Fig. 3.5 (a) Three-port model of complete transducer.

(b) Equivalent circuit as seen from electrical terminals.

26 Fig. 3.6 Input admittance of a uniform interdigital transducer, Eq. 3.6, as a function of normalized frequency.

32 Fig. 3.7 First-order reflections from the edges of the transducer fingers.

33 Fig. 3.8 Equivalent circuit of one section of a transducer including mismatch at finger edges. $\theta_i' = 2\pi a_i / \lambda'$, $\theta = \pi (L_i - a_i) / \lambda$, $\lambda = v/f$, $\lambda' = v'/f$, $v' = v / (1 + K^2/2)$ and the resonant frequency $f_i = \frac{1}{2} \left[\frac{a_i}{v} + \frac{L_i - a_i}{v'} \right]^{-1}$.

34 Fig. 3.9 (a) Geometry of a split-electrode transducer.

(b) Reflections from finger edges of a split-electrode transducer.

37 Fig. 3.10 (a) Schematic operation of a three-phase unidirectional transducer.

(b) Geometry of a three-phase unidirectional transducer.

39 Fig. 3.11 Relative amplitude of excitation of fundamental and harmonics of a uniform transducer as a function of metallization ratio, normalized to fundamental value at $a/L = 0.5$ (4).

41 Fig. 3.12 Transducer efficiency in decibels as a function of aperture for different sheet resistivities.

BIBLIOGRAPHY

A. Review Articles and Proceedings of Special Conferences

G1. Auld, B.A. "Acoustic Fields and Waves in Solids," John Wiley, (1973), Vols. 1 and 2.

G2. Farnell, G.W. "Properties of Elastic Surface Waves," in Physical Acoustics, W. Mason and R. Thurston, (Eds.), Academic Press, 6, (1970), 109-166.

G3. Farnell, G.W., and Adler, E.L. "Elastic Wave Propagation in Thin Layers," in Physical Acoustics, W. Mason and R.N. Thurston, (Eds.), Academic Press, 9, (1972), 35-127.

G4. IEE Conference Publication No. 109, Aviemore, Scotland, (1973). "Component Performance and Systems Applications of Surface Acoustic Wave Devices."

G5. IEEE Ultrasonics Symposium Proc., Boston, Massachusetts, (1972).

G6. IEEE Ultrasonics Symposium Proc., Monterey, California, (1973).

G7. IEEE Trans. Microwave Theory Tech., MTT-17, (1969), 798-1052. "Special issue on Microwave Acoustics."

G8. IEEE Trans. Sonics Ultrasonics, SU-20, (1973), 79-230.

Also

IEEE Trans. Microwave Theory Tech., MTT-21, (1973), 161-312. "Special issue on Microwave Acoustic Signal Processing."

G9. Kino, G.S.,and Matthews, H. "Signal Processing in Acoustic Surface Wave Devices," IEEE Spectrum, 8, (1971), 22-35.

G10. Maines, J.D., and Paige, E.G.S. "Surface-acoustic-wave Components, Devices and Applications," Proc. IEE, Reviews, 120, (1973), 1078-1110.

G11. White, R.M. "Surface Elastic Waves," IEEE Proc., 58, (1970), 1238-1276.

B. Alphabetic Listing

1. Adler, E.L., Bridoux, E., Coussot, G., and Dieulesaint, E. "Harmonic Generation of Acoustic Surface Waves in B_{12} Ge 0_{20} and Li Nb 0_3," IEEE Trans. Sonics Ultrasonics, SU-20, (1973), 13-16.

2. Adler, R. "Surface-Wave Amplifier with Improved Geometry," Electron. Lett., 8, (1972), 65-66.

3. Angerstein, J., and Voges, E. "Photoresist Techniques for Acoustic Surface Wave Transducers," AEU, 26, (1972), 387-390.

4. Bahr, A.J., Lee, R.E., Cristal, E.G., and Gysel, U. H. "Microwave Network Synthesis," ECOM Report No. 0127-F, (1973).

5. Bell, D.T. Jr., Holmes, J.D., and Ridings, R.V. "Application of Acoustic Surface-wave Technology to Spread Spectrum Communications," IEEE Trans. Sonics Ultrasonics, SU-20, (1973), 181-189.

6. Blotekjaer, K., Ingebrigtsen, K.,and Skeie, H. "A Method for Analysing Waves in Structures Consisting of Metal Strips on Dispersive Media," IEEE Trans. Electron Devices, ED-20, (1973), 1133-1138.

7. Blotekjaer, K., Ingebrigtsen, K.,and Skeie, H. "Acoustic Surface Waves in Piezoelectric Materials with Periodic Metal Strips on the Surface," IEEE Trans. Electron Devices, ED-20, (1973), 1139-1146.

8. Broers, A.N., and Hatzakis, M. "Electron Beam Fabrication on High Frequency, Low Loss, Wide Band-width Interdigital Acoustic Surface Wave Transducers," AFCRL Report No. 71-0323, (1971).

9. Campbell, J.J., and Jones, W.R. "A Method for Estimating Optimal Crystal Cuts and Propagation Directions for Excitation of Piezoelectric Surface Waves," IEEE Trans. Sonics Ultrasonics, SU-15, (1968), 209-217.

10. Carr, P.H., De Vito, P., and Szabo, T.L. "The Effect of Temperature and Doppler Shift on the Performance of Elastic Surface Wave Encoders and Decoders," IEEE Trans. Sonics Ultrasonics, SU-19, (1972), 357-367.

11. Cheek, T.F., Hays, R.M., and Hartmann, C.S. "A Wide-Band Low-Shape-Factor Amplifier Module Using an Acoustic Surface-wave Bandpass Filter, IEEE J. of Solid State Circuits, SC-8, (1973), 66-70.

12. Claiborne , L.T., Staples, E.J., and Harris, J.L. "MOSFET Ultrasonic Surface Wave Detectors for Programmable Matched Filters," Appl. Phys. Lett., 19, (1971), 58-60.

13. Coldren, L.A., and Kino, G.S. "Monolithic Acoustic-Surface-Wave Amplifier," Appl. Phys. Lett., 18, (1971), 317-319.

14. Davies, L.W., and Lawrence, M.W. "Prospects for Surface Elastic Wave Crystal-controlled Delay Line Oscillators," Proc. IREE (Australia), 32, (1971), 61-62.

15. De Vito, P.A., Carr, P.H., Kearns, W.J., and Silva, J.H. "Encoding and Decoding with Elastic Surface Waves at 10 Megabits per Second," IEEE Proc., 59, (1971), 1523-1525.

16. De Vries, A.J., Dias, J.F., Rypkema, J.N., and Wojcik, T.J. "Characteristics of Surface-Wave Integrable Filters (SWIFS)," IEEE Trans. Broadcast Telev. Rec., BTR-17, (1971), 16-22.

17. Duffy, M.T., Wang, C.C., O'Clock, G.D., McFarlane, III, S.H., and Zanzucchi, P.J. "Epitaxial Growth and Piezoelectric Properties of Al N, GaN and GaAs on Sapphire or Spinel," J. of Electronic Materials, 2, (1973), 359-372.

18. Emtage, P.R. "Description of Interdigital Transducers," J. Appl. Phys., 43, (1972), 4486-4489.

19. Engan, H. "Excitation of Elastic Surface Waves by Spatial Harmonics of Interdigital Transducers," IEEE Trans. Electron Devices, ED-16, (1969), 1014-1017.

20. Farnell, G.W., Cermak, I.A., Silvester, P., and Wong, S.K. "Capacitance and Field Distributions for Interdigital Surface-Wave Transducers," IEEE Trans. Sonics Ultrasonics, SU-17, (1970), 188-195.

21. Gerard, H.M., Judd, G.W., and Pedinoff, M.E. "Phase Corrections for Weighted Acoustic Surface Wave Dispersive Filters," IEEE Trans. Microwave Theory Tech., MTT-20, (1972), 188-192.

22. Guyot, Joel. "Analyse Spectrale au Moyen de Filtres Dispersifs," Thése de Docteur-Ingenieur, Faculté des Sciences, Nice, France, (1973).

23. Hagon, P.J., Lakin, K.M., and Seymour, R.N. "Programmable Acoustic Signal Processing Devices," ECOM Report No. 05011-1, (1974).

24. Hagon, P.J., Micheletti, F.B., Seymour, R.N., and Wrigley, C.Y. "A Programmable Surface Acoustic Wave Matched Filter for Phase-Coded Spread Spectrum Waveforms," IEEE Trans. Sonics Ultrasonics, SU-20, (1973), 221-224.

25. Hanebrekke , H., and Ingebrigtsen, K.A. "Acoustoelectric Amplification of Surface Waves in Structure of Cadmium-Selenide Film on Lithium Niobate," Electron. Lett., 6, (1970), 520-521.

26. Hartmann, C.S., Bell, D.T., and Rosenfeld, R.C. "Impulse Model Design of Acoustic Surface-Wave Filters," IEEE Trans. Sonics Ultrasonics, SU-20 (1973) 80-93.

27. Hartmann, C.S., Jones, W.S., and Vollers, H. "Wideband Unidirectional Interdigital Surface Wave Transducers," IEEE Trans. Sonics Ultrasonics, SU-19, (1972) 378-381.

28. Hickernell, F.S. "dc Triode Sputtered Zinc Oxide Surface Elastic Wave Transducers," J. Appl. Phys., 44, (1973), 1061-1071.

29. Hughes, A.J. "Elastic Surface Wave Guidance by $\Delta v / v$ Effect Guidance Structures," J. Appl. Phys., 43, (1972), 2569-2586.

30. Hunsinger, B.J. "Surface Acoustic Wave Devices and Applications, 3. Spread Spectrum Processors," Ultrasonics, 11, (1973), 254-262.

31. Hunsinger, B.J., and Franck, A.R. "Programmable Surface Wave Tapped Delay Line," IEEE Trans. Sonics Ultrasonics, SU-18, (1971), 152-154.

32. Hurlburt, D.H. "Philosophy and Design of Acoustic Surface Wave Filters," to appear in RCA Engineer, 19, (1974).

33. Ingebrigtsen, K.A. "Surface Waves in Piezoelectrics," J. Appl. Phys., 40, (1969), 2681-2686.

34. Javed, A. "Effect of Surface States on Surface-Wave Amplification in a Composite Structure of CdSe Film on Li Nb 0_3," IEEE Trans. Electron Devices, ED-19, (1972), 1296-1297.

35. Jones, W.S., Hartmann, C.S., and Claiborne, L.T. "Evaluation of Digitally Coded Acoustic Surface-Wave Matched Filters," IEEE Trans. Sonics Ultrasonics, SU-18, (1971), 21-27.

36. Jones, W.S., Hartmann, C.S., and Strundivant, T.D. "Second Order Effects in Surface Wave Devices," IEEE Trans. Sonics Ultrasonics, SU-19, (1972), 368-377.

37. Judd, G., Morse, F., and Smith, W.R. Acoustic Signal Processing Devices," ECOM Report No. 0023-1, (1972), and Report No. 0023-F, (1973).

38. Kharusi, M.S. and Farnell, G.W. "Diffraction and Beam Steering for Surface-Wave Comb Structures on Anisotropic Substrates," IEEE Trans. Sonics Ultrasonics, SU-18, (1971), 35-42.

39. Kino, G.S. "F.E.T. Theory of Strip Coupled Rayleigh Wave Amplifier," Elec. Lett., 9, (1973), 52-53.

40. Kino, G.S., and Coldren, L.A. "Noise Figure Calculation for the Rayleigh Wave Amplifier," Appl. Phys. Lett., 22, (1973), 50-52.

41. Kino, G.S., and Gautier, H. "Convolution and Parametric Interaction with Semiconductors," J. Appl. Phys., 44, (1973), 5219-5221.

42. Kino, G.S., Ludvik, S., Shaw, H.J., Shreve, W.R., White, J.M., and Winslow, D.K. "Signal Processing by Parametric Interactions in Delay-Line Devices," IEEE Trans. Microwave Theory Tech., MTT-21, (1973), 244-255.

43. Kino, G.S., and Reeder, T.M. "A Normal Mode Theory for the Rayleigh Wave Amplifier," IEEE Trans. Electron Devices, ED-18, (1971), 909-919.

44. Kodak Publication No. P-79. "An Introduction to Photofabrication Using Kodak Photosensitive Resists."

45. Lagasse, P.E., Mason, I.M., and Ash, E.A. "Acoustic Surface Waveguides, Analysis and Assessment," IEEE Trans., SU-20, (1973), 143-153.

46. Lakin, K.M. "Electrode Resistance Effects in Interdigital Transducers," IEEE, MTT-22, (1974), 418-424.

47. Lakin, K.M., "Perturbation Theory for Electromagnetic Coupling to Elastic Surface Waves on Piezoelectric Substrates," J. Appl. Phys., 42, (1971), 899-906.

48. Lakin, K.M., and Shaw, H.J. "Surface Wave Delay Line Amplifiers," IEEE Trans. Microwave Theory Tech., MTT -17, (1969), 912-920.

49. Lean, E.G., Powell, C.G., and Kuhn, L. "Acoustic Surface Wave Mixing on α - Quartz," Appl. Phys. Lett., 15, (1969), 10-12.

50. Lean, E.G., and Tseng, C.C. "Nonlinear Effects in Surface Acoustic Waves," J. Appl. Phys., 41, (1970), 3912-3917.

51. Lewis, M.F. "Surface-Acoustic Wave Filters Employing Symmetric Phase Weighted Transducers," Elec. Lett., 9, (1973), 138-140.

52. Lim, T.C., Kraut, E.A.,and Thompson, R.B. "Nonlinear Materials for Acoustic Surface Wave Convolver," Appl. Phys. Lett., 20,(1972), 127-129.

53. Lopen, P.O. "Second-Harmonic Generation in an Elastic Surface Wave in α - Quartz," J. Appl. Phys., 39, (1968), 5400-5404.

54. Ludvik, S., and Quate, C.F. "Amplification of Surface Shear Wave Mode in Ga As," J. Appl. Phys., 43, (1972), 3619-3622.

55. Luukkala, M., and Surakka, J. "Acoustic Convolution and Correlation and the Associated Nonlinearity Parameters in Li Nb 0_3," J. Appl. Phys., 43, (1972), 2510-2518.

56. Maerfeld, C., and Tournois, P. "Perturbation Theory for the Surface Wave Multistrip Coupler," Elec. Lett., 9, (1973), 115-116.

57. Marshall, F.G., and Paige, E.G. "Novel Acoustic Surface Wave Directional Coupler with Diverse Applications," Elec. Lett., 7, (1971), 460-462.

58. Marshall, F.G., and Paige, E.G.S. "Observed Properties of an Acoustic Surface Wave Multistrip Coupler," Elec. Lett., 7, (1971), 463-464.

59. Marshall, F.G., Newton, C.O., and Paige, E.G.S. "Theory and Design of the Surface Acoustic Wave Multistrip Coupler," IEEE Trans. Sonics Ultrasonics, SU-20, (1973), 124-133.

60. Marshall, F.G., Newton, C.O., and Paige, E.G.S. Surface Acoustic Wave Multistrip Components and Their Applications," IEEE Trans. Sonics Ultrasonics, SU-20, (1973), 134-142.

61. Mason, I.M. "Anisotropy, Diffraction Scaling, Surface Wave Lenses, and Focussing," JASA, 53, (1973), 1123-1128.

62. Mason, I.M., and Ash, E.A. "Acoustic Surface-Wave Beam Diffraction on Anisotropic Substrates," J. Appl. Phys., 42, (1971), 5343-5351.

63. Matthaei, G.L. "Acoustic Surface Wave Transversal Filters," IEEE Trans. Circuit Theory, CT-20, (1973), 459-470.

64. McClellan, J.H., Parks, T.W., and Rabiner, L.R. "A Computer Program for Designing Optimum FIR Linear Phase Digital Filters," IEEE Trans. AU-21, (1973), 506-526.

65. Meyer, P.C., and Schultz, M.B. "Temperature Effects in Reflective Surface Acoustic Wave Delay Lines," Elec. Lett., 9, (1973), 523-525.

66. Munasinghe, M., and Farnell, G.W. "Acoustic Surface Wave Scattering on a Homogeneous Three-Quarter Space," J. Appl. Phys., 44, (1973), 2025-2031.

67. Nakagawa, Y., Yamanouchi, K., and Shibayama, K. "Third Order Elastic Constants of Lithium Niobate," J. Appl. Phys., 44, (1973), 3969-3974.

68. O'Clock, G.D., and Duffy, M.T. "Acoustic Surface Wave Properties of Epitaxially Grown Aluminum Nitride and Gallium Nitride on Sapphire," Appl. Phys. Lett., 23, (1973), 55-56.

69. O'Clock, G.D., Grasse, C.L., and Gandolfo, D.A. "Switchable Acoustic Surface Wave Sequence Generator," IEEE Proc. 59, (1971), 1536-1537.

70. Oliner, A.A., Bertoni, H.L.,and Li, R.C. "Microwave Network Formalism for Acoustic Waves in Isotropic Media," Proc. IEEE, 60, (1972), 1503-1517.

71. Otto, O.W. "Real Time Fourier Transform with a Surface-Wave Convolver," Elec. Lett., 8, (1972), 622-623.

72. Otto, O.W., and Moll, N.J. "Lithium-Niobate-Silicon Surface Wave Convolver," Elec. Lett., 8 (1972), 600-602.

73. Ozdemir, F.S., Wolf, E.D., and Buckey, C.R. "Computer-Controlled Scanning Electron Microscope System for High Resolution Microelectronic Pattern Fabrication," IEEE Trans. Electron Devices, ED-19, (1972) 624-628.

74. Schulz, M.B., and Holland, M.G. "Surface Acoustic Wave Delay Lines with Small Temperature Coefficient," IEEE Proc. Lett., 58, (1970), 1361-1362.

75. Schulz, M.B., and Matsinger, J.H. "Rayleigh-Wave Electromechanical Coupling Constants," Appl. Phys. Lett., 20, (1972), 367-369.

76. Schulz, M.B., Matsinger, B.J., and Holland, M.G. "Temperature Dependence of Surface Acoustic Wave Velocity on α-Quartz," J. Appl. Phys., 41, (1970), 2755-2765.

77. Shaw, H.J. "Surface Wave Parametric Processors and Long Delay Lines," ECOM Report 0-134S, December 1973, (Stanford).

78. Shuskus, A.J., Reeder, T.M., and Paradis, E.L. "RF-Sputtered Aluminum Nitride Films on Sapphire," Appl. Phys. Lett., 24, (1974), 155-156.

79. Slobodnik, A.J.,Jr. "A Review of Material Tradeoffs in the Design of Acoustic Surface Wave Devices at VHF and Microwave Frequencies," IEEE Trans. Sonics Ultrasonics, SU-20, (1973), 315-323.

80. Slobodnik, A.J., Jr. "Nonlinear Effects in Microwave Acoustic Li Nb 0_3 Surface-Wave Delay Lines," J. Acoust. Soc. Am., 48, (1970), 203-210.

81. Slobodnik, A.J., Jr. "UHF and Microwave Frequency Acoustic Surface Wave Delay Lines: Design," Air Force Cambridge Research Lab., Report No. AFCRL-TR-73-0538, (1973).

82. Slobodnik, A.J., Jr., and Conway, E.D. "New High-Frequency High-Coupling Low-Beam-Steering Cut for Acoustic Surface Waves on Li Nb 0_3," Elec. Lett., 6, (1970), 171-173.

83. Slobodnik, A.J., Jr., and Szabo, T.L. "Minimal Diffraction Cuts for Acoustic-Surface Wave Propagation on Bi_{12} Ge 0_{20}," J. Appl. Phys., 44, (1973), 2937-2941.

84. Smith, H.I., Bachner, F.J.,and Efremow, N. "A High-Yield Photolithographic Technique for Surface Wave Devices," J. Electrochem. Soc., 118, (1971), 821-825.

85. Smith, W.R., Gerard, H.M., Collins, J.H., Reeder, T.M., and Shaw, H.J. "Analysis of Interdigital Surface Wave Transducer by use of an Equivalent Circuit Model," IEEE Trans. Microwave Theory Tech., MTT-17, (1969), 856-864.

86. Smith, W.R., Gerard, H.M., Collins, J.H., Reeder, T.M., and Shaw, H.J. "Design of Surface Wave Delay Lines with Interdigital Transducers," IEEE Trans. Microwave Theory Tech., MTT-17, (1969), 865-873.

87. Smith, W.R., Gerard, H.M., and Jones, W.R. "Analysis and Design of Dispersive Interdigital Surface-Wave Transducers," IEEE Trans. Microwave Theory Tech., MTT-20, (1972), 458-471.

88. Smith, W.R., Gerard, H., and Snow, P.B. "Highly Dispersive Acoustic Filter Study," ECOM Report No. 0046-F, (1973).

89. Spears, D.L., and Smith, H.I. "X-ray Lithography - A New High Resolution Replication Process," Solid State Technology, 15, (1972), 21-26.

90. Squire, W.D., Whitehouse, H.J., and Alsup, J.M. "Linear Signal Processing and Ultrasonic Transversal Filters," IEEE Trans. Microwave Theory Tech., MTT-17, (1969), 1020-1040.

91. Staples, E.J., and Claiborne, L.T. "A Review of Device Technology for Programmable Surface Wave Filters," IEEE Trans. Sonics Ultrasonics, SU-20, (1973), 197-205.

92. Stern, E. "Microsound Components, Circuits and Applications," IEEE Trans. Microwave Theory Tech., MTT-17, (1969), 835-844.

93. Szabo, T.L., and Slobodnik, A.J., Jr. "Acoustic Surface Wave Diffraction and Beam Steering," Air Force Cambridge Research Lab Report AFCRL-TR-73-0302, (1973).

94. Szabo, T.L., and Slobodnik, A.J., Jr. "The Effect of Diffraction on the Design of Acoustic Surface Wave Devices," IEEE Trans. Sonics Ultrasonics, SU-20, (1973), 240-251.

95. Tancrell, R.H. "Analytic Design of Surface Wave Bandpass Filters," IEEE Trans. Sonics Ultrasonics, SU-21, (1974), 12-22.

96. Tancrell, R.H., and Holland, M.G. "Acoustic Surface Wave Filters," IEEE Proc. 59, (1971), 393-409.

97. Tancrell, R.H., and Sandy, F. "Analysis of Interdigital Transducers for Acoustic Surface Wave Devices," Air Force Cambridge Research Lab Report No. AFCRL-TR-73-0030, (1973).

98. Tiersten, H.F. "Elastic Surface Waves Guided by Thin Films," J. Appl. Phys., 40, (1969), 770-789.

99. Tseng, C.C. "Frequency Response of an Interdigital Transducer for Excitation of Surface Elastic Waves," IEEE Trans. Electron Devices, ED-15, (1968), 586-594.

100. Tu, C.C., and Farnell, G.W. "Thickness Effects in Overlay Guides for Elastic Surface Waves," IEEE Trans. Sonics Ultrasonics, SU-20, (1973), 302-314.

101. Vella, P.J., and Stegeman, G.I. "Surface Wave Harmonic Generation on Sapphire and α - quartz," Appl. Phys. Lett., 23 (1973), 505-507.

102. Wauk, M.T., and Winslow, D.K. "Vacuum Deposition of Al N Acoustic Transducers," Appl. Phys. Lett., 13, (1968), 286-288.

103. White, R.M. "Surface Elastic-Wave Propagation and Amplification," IEEE Trans. Electron Devices, ED-14, (1967), 181-189.

104. White, R.M., and Voltmer, F.W. "Direct Piezoelectric Coupling to Surface Elastic Waves," Appl. Phys. Lett., 7, (1965), 314-316.

105. White, R.M., and Voltmer, F.W. "Ultrasonic Surface Wave Amplification in Cadmium Sulfide," Appl. Phys. Lett., 8, (1966), 40-42.

106. Williamson, R.C., and Smith, H.I. "The Use of Surface Elastic Wave Reflection Gratings in Large Time Bandwidth Pulse Compression Filters," IEEE Trans. Sonics Ultrasonics, SU-20, (1973), 113-123.

FARNELL, G.W.
--An overview of acoustic surface-wave technology.