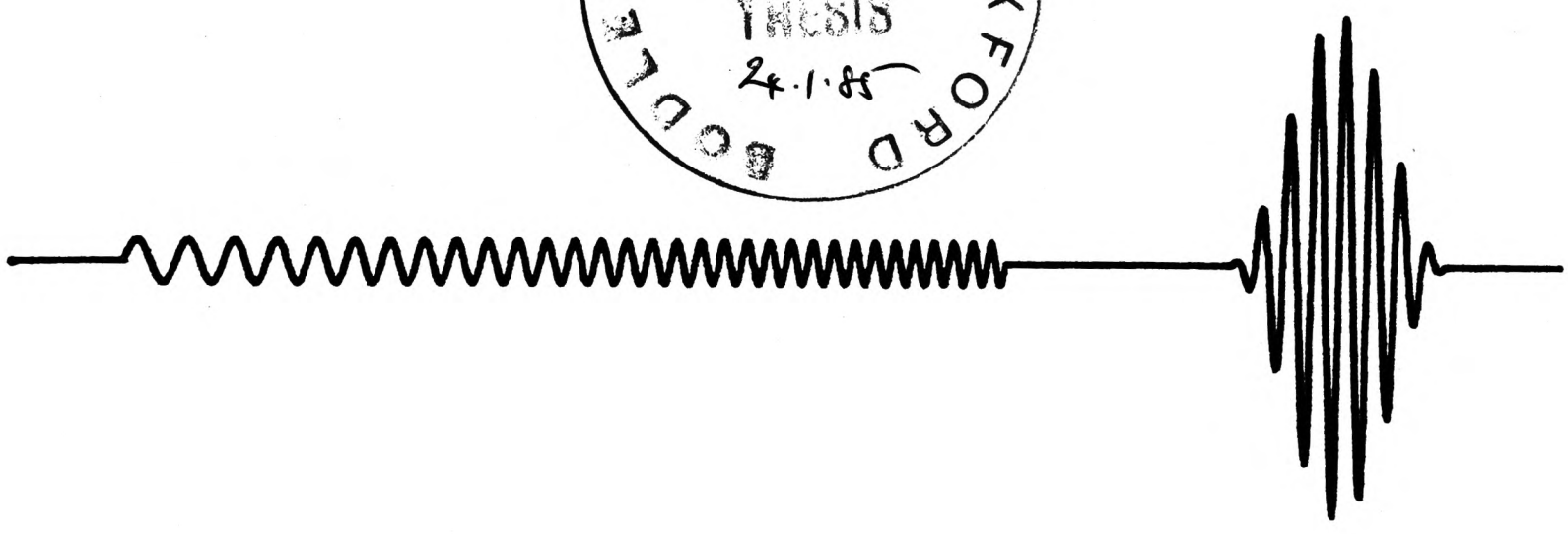
THE INTERACTION OF SURFACE ACOUSTIC WAVES WITH ARRAYS OF THIN METAL DOTS

A Thesis for the Degree of
Doctor of Philosophy
by
Frederick Huang
of
University College, Oxford
Hilary Term, 1984

Supervised by
Prof. E. G. S. Paige, FRS

BODLEIAN LIBRARY OXFORD
DEPOSITED
THESIS
24.1.85



"His Voice as the Sound of Many Waters" (Rev. 1:15)

ABSTRACT

The Interaction of Surface Acoustic Waves with Arrays of Thin Metal Dots

A Thesis for the degree of Doctor of Philosophy by Frederick Huang of University College, Oxford, submitted during Hilary Term, 1984.

One class of Surface Acoustic Wave (SAW) devices employs reflecting arrays to determine device characteristics. Dots of metal film and grooves have been used as reflectors; reflection is caused by the mechanical discontinuity. This research project is a study of an alternative reflector, the thin metal dot, which reflects by interacting with the electric field that accompanies the SAW on a piezo-electric material.

Experimental work is done on Y-cut $LiNbO_3$, a common SAW material, the aim being to characterise the reflection mechanism sufficiently to design devices, particularly on this material. Reflectivities of a large range of dot sizes is reported. Velocity change due to dots, another parameter required in device design, is also measured. In addition, holes in a metal film, physically the opposite of dots, are also studied.

A theorectical model is developed to predict the behaviour of dots in devices. For a Z-propagating wave in $LiNbO_3$, the SAW velocity for completely free and completely metallised surfaces were used to specify the properties of the material. For propagation in other directions on the surface, particularly near the X direction, a different acoustic mode, the Leaky Surface Wave (LSW), has an influence on SAW reflections, and the model also requires the LSW velocities under the same two conditions. The model uses these parameters, together with dot geometry, to determine reflectivity in any direction and velocity change due to the dots. The results agree well with experiments.

A device (a RAC) is designed using this model, to bring out any unforseen difficulties in the use of dot arrays. The device performed as expected except for some unanticipated attenuation in the arrays, which can almost certainly be predictably allowed for in future designs, being broadband and not very severe.

ACKNOWLEDGEMENT

Much gratitude is due to my supervisor, Ted Paige, for first kindling my interest in Surface Acoustic Waves, and for subsequent guidance, encouragement and opportunities. The other members of the SAW group have also been very helpful. In particular, the software for those experiments which are computer controlled was developed by them. There has not been a need for any computer controlled experiment not already in existence.

The experimental work has involved a large number of devices. These were fabricated by Mr. Brian French at Oxford, and by Roger Gibbs via an arrangement with Plessey Research, Caswell.

The project was financed by the SERC. Personal financial support was provided by Racal Research Ltd. and the Department of Education and Science.

CONTENTS

Abstract 2

Acknowledgement 3

1. INTRODUCTION 1- 1
   1.1 Surface Acoustic Waves and SAW Devices 1- 1
   1.2 The Aim of this Project 1- 6

2. EXPERIMENTAL DETERMINATION OF DOT PROPERTIES 2- 1
   2.1 Reflectivity (Z to X) of Widely Separated Dots 2- 3
      2.1.1 Layout of devices 2- 3
      2.1.2 Measurements 2- 9
      2.1.3 Calculations 2-10
   2.2 180 Degree Reflections by Dots 2-11
   2.3 Further Experiments on Reflectivity (Z-X), Including Dots with Near Neighbours 2-14
   2.4 Reflection by Holes 2-15
   2.5 Velocity Change in Dot Arrays: Z direction .. 2-16
      2.5.1 Device Layout 2-16
      2.5.2 Measurements 2-17
      2.5.3 Calculations 2-19
   2.6 Velocity Change in Dot Arrays: X Direction .. 2-24
   2.7 Further Experiments on Velocity and 180 Degree Reflections 2-25
      2.7.1 Device Layout 2-25
      2.7.2 Measurements 2-25
      2.7.3 Calculations 2-27
   2.8 Velocity of the Leaky Surface Wave 2-29
   2.9 Radiation Resistance of Transducers in Various Directions 2-32

3. THEORY 3- 1
   3.1 The Generation of SAWs by Electric Charge Varying in One Dimension 3- 2
   3.2 The Effect of the LSW 3- 6
   3.3 Reflections and Scattering by Widely-Separated Dots 3- 8
      3.3.1 Charge on a dot 3- 8
      3.3.2 Reflections through 90 Degrees 3-12
      3.3.3 Reflections in Other Directions 3-16
      3.3.4 Anisotropy 3-17
      3.3.5 Further Computations 3-18
   3.4 Reflections and Scattering by Dots with Near-Neighbours 3-19
   3.5 Velocity Change 3-23
   3.6 90 Degree Reflections by Holes 3-25
   3.7 Simplified Version of Scattering Theory 3-27
   3.8 Summary 3-33

4. RESULTS 4- 1
   4.1 The evaluation of Ge 4- 1
   4.2 Comparison of Ge with dot data 4- 2
   4.3 Reflections through 90 Degrees and Scattering by Widely-Separated Dots 4- 4
   4.4 Additional Results on Reflectivity, Including Dots with Interactions 4-14
   4.5 Velocity Change Due to Dots 4-14

4.6 Reflections by Holes 4-20
4.7 Plots of Transducer Radiation Resistance 4-24
4.8 Reflections by Strips through 180 Degrees ... 4-26
4.9 Reflections through 90 Degrees
by Strips Inclined at 45 Degrees 4-27
4.10 Reflections from Circular Dots 4-31

5. THE DESIGN OF A PULSE COMPRESSOR 5- 1
5.1 The Principle of Stationary Phase 5- 2
5.2 Chirped Signals and their Matched Filters ... 5- 3
5.3 Standard RAC Equations 5- 8
5.4 Considerations in the Design of a Dot RAC ... 5-10
5.5 Features of the Selected Array Pattern 5-14
5.5.1 Scattering of SAWs by the Array 5-14
5.5.2 Scatttering of Bulk and LSWs 5-18
5.6 Data on Dots and Transducers used in the
RAC Design 5-19
5.7 Computer Calculations for the RAC 5-25

6.PERFORMANCE OF THE PULSE COMPRESSOR 6- 1
6.1 Temperature Control of the Devices 6- 1
6.2 Measurement of Frequency Response 6- 2
6.3 Results and Observations 6- 6
6.3.1 Amplitude of Frequency Response 6- 6
6.3.2 Phase of Frequency Response 6- 9
6.4 Summary of Observations 6-16

7. CONCLUSIONS 7- 1
7.1 The Outcome of this Project 7- 1
7.2 Possibilities for Further Development 7- 3

APPENDICES:
A: The Insertion-loss Measurement System 8- 1
B: The Equipment for Measuring Phase Response
of a RAC 8- 6
C: 180 Degree Reflections by Strips at
Normal Incidence 8- 9
D: 90 Degree Reflections from Strips 8-13
E: Frequency Respose of an Array in L-Path 8-15
F: Attenuation of an Incident Wave through
Reflections 8-18
G: Power Lost By Reflections in a Chirped Array; .. 8-21
Outline of Calculations to Find Power Reflected 8-23
H: The Effect of Position Errors in a RAC
on its Response 8-27
I: Waves Reflected by a Distribution of Scatterers; 8-30
The Amplitude of the Scattered Wave; 8-30
Power Lost by a Randomised Array 8-31

REFERENCES 8-34

CHAPTER 1

INTRODUCTION

1.1 Surface Acoustic Waves and SAW Devices

Surface Acoustic Wave (SAW) devices are now used for signal processing in many electronic systems ranging from domestic television receivers to radar. In a SAW device, the electrical signal is first converted by a transducer into a high frequency (tens or hundreds of MHz) acoustic wave on the surface of a solid, and eventually converted back to an electrical signal by a second transducer. The desired processing, such as frequency filtering, is achieved either by suitable design of the transducers or by placing frequency-selective arrays of reflecting elements in the path of the acoustic beam.

In introducing SAWs, it is useful to first refer to Bulk Waves in solids, which are fairly well known. They are similar to sound waves in air except that a solid can support transverse as well as longitudinal waves. It was shown theoretically in 1885 by Lord Rayleigh that an elastic or acoustic wave trapped near the surface of the solid is also possible. (It is not necessary to have a layer of a different

material.) Most of the energy in this wave is within a wavelength of the surface. This is probably the most useful feature of the surface wave: reflectors and transducers only need to be placed on the surface; the wave is very accessible. This planar arrangement of the device also simplifies fabrication. Much of the technology developed for integrated circuits is also suitable for SAW devices. A further advantage is that the thickness of the substrate material used has little effect on the wave. This is because almost all the energy is within a few microns or tens of microns from the surface irrespective of the thickness of the material, which therefore can be made as thick as is structurally required. A drawback is that SAW devices are very sensitive to surface contamination.

Almost all substrate materials used in SAW devices are piezo-electric. A varying electric field accompanies the periodic stresses of the wave, enabling it to be coupled to external electrical circuits. The electric field can be shorted out at the surface by a metallic film, reducing the wave velocity. The fractional change in velocity $\frac{\Delta v}{v}$ between the plated and unplated substrate is a very useful parameter indicating the strength of piezoelectric coupling. It is often possible to determine device behaviour solely in terms of $\frac{\Delta v}{v}$ and transducer or reflector geometry.

The velocity of the SAW is typically a few thousand metres per second, five orders of magnitude less than electromagnetic waves in free space. On a substrate of typical size, the saw path-length can be several thousand

wavelengths long; this feature is obviously useful in devices such as delay lines and narrowband filters (which have long impulse lengths).

Further device details are most easily described in an example. A SAW device with demanding specifications is the radar pulse compressor, sometimes implemented using reflecting arrays. The Reflecting-Array Compressor (RAC) is shown in fig. 1.1, and the chirped signal to which it is matched in fig. 1.2 . In this example, the instantaneous frequency of the signal increases with time. The individual reflectors of the RAC are weakly reflecting so that significant reflection occurs only when there is constructive interference between a large number of reflectors. When an echo signal with the waveform shown in fig. 1.2 is applied to the RAC, the high frequencies are reflected at P, and the lower frequencies at Q; the lower frequency energy which arrives earlier is delayed for a greater time and the pulse is thus compressed by up to about 10 000 : 1 in some devices [11]. The pulse produced by the transmitter can thus be spread out in time, reducing peak power requirements. The compression ratio is proportional to TB, where T is the duration of the uncompressed pulse, and B its bandwidth. A more mathematical description is defered to a later chapter.

The reflectors are grooves milled into the substrate, strips of metal film, or rows of metal-film dots on the surface. In all cases, the wave is reflected by the step discontinuity in the surface ('topographical' reflections). If the dots or strips are of a dense material such as gold,

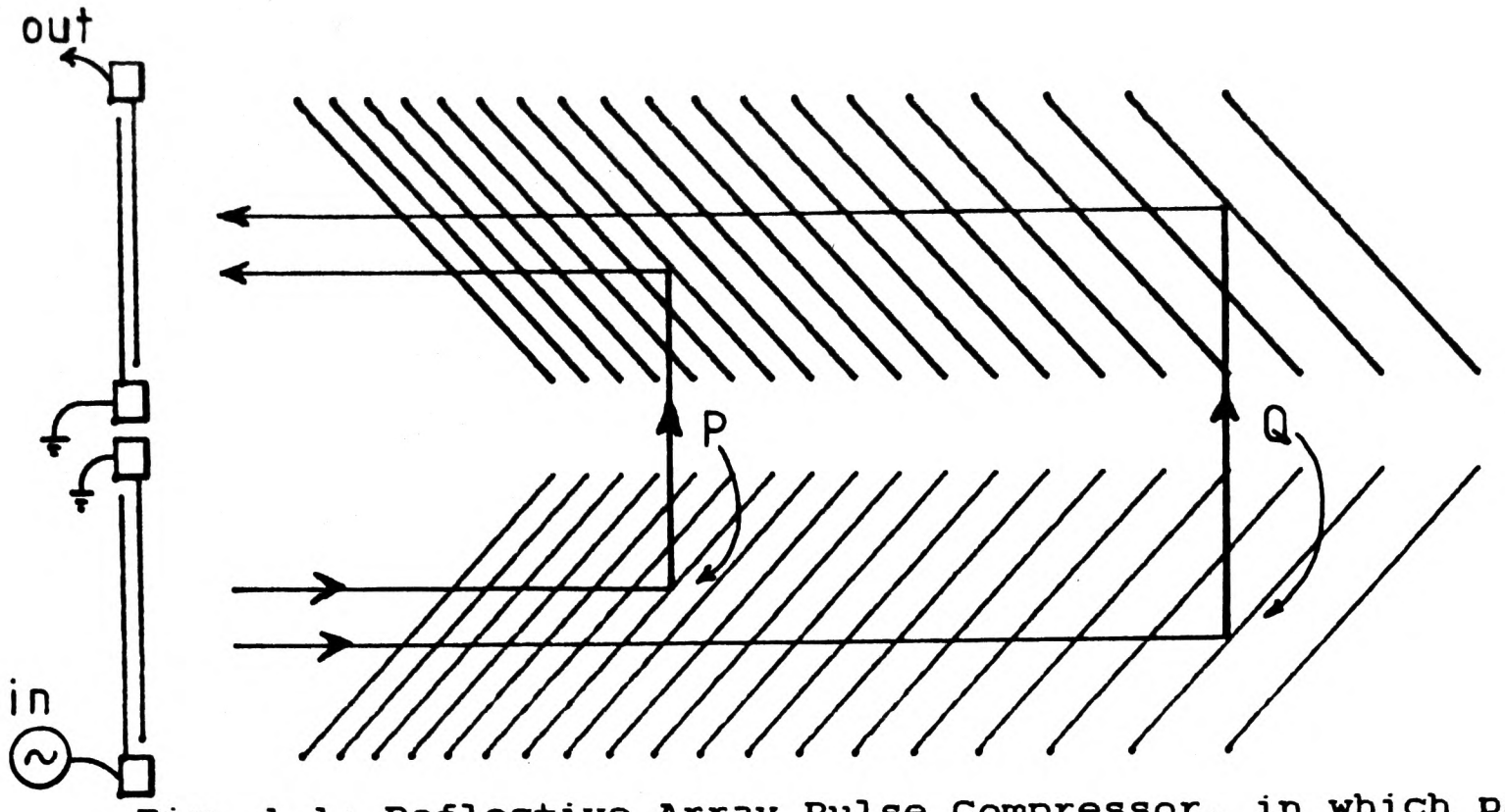


Fig. 1.1: Reflective Array Pulse Compressor, in which path length depends on frequency.

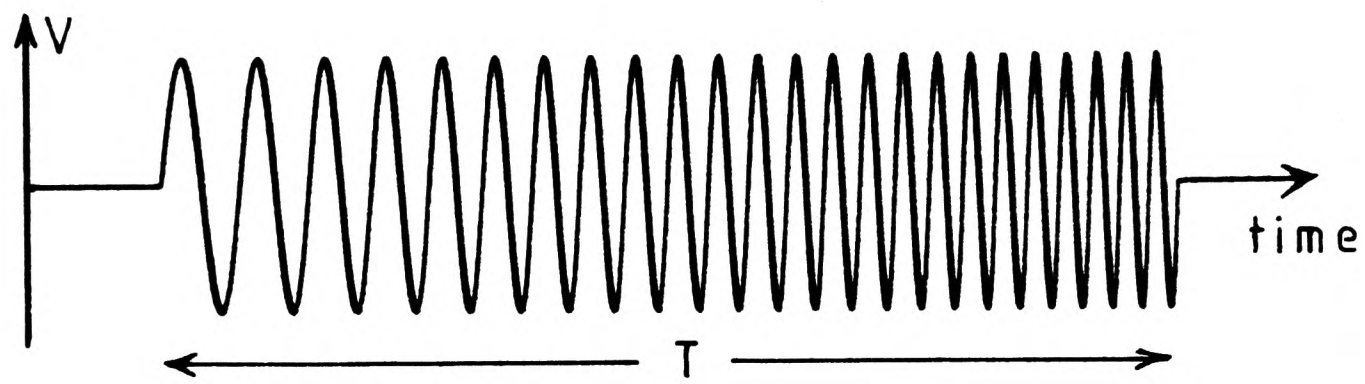


Fig. 1.2: Chirped signal, with varying instantaneous frequency, to which the above filter is matched.

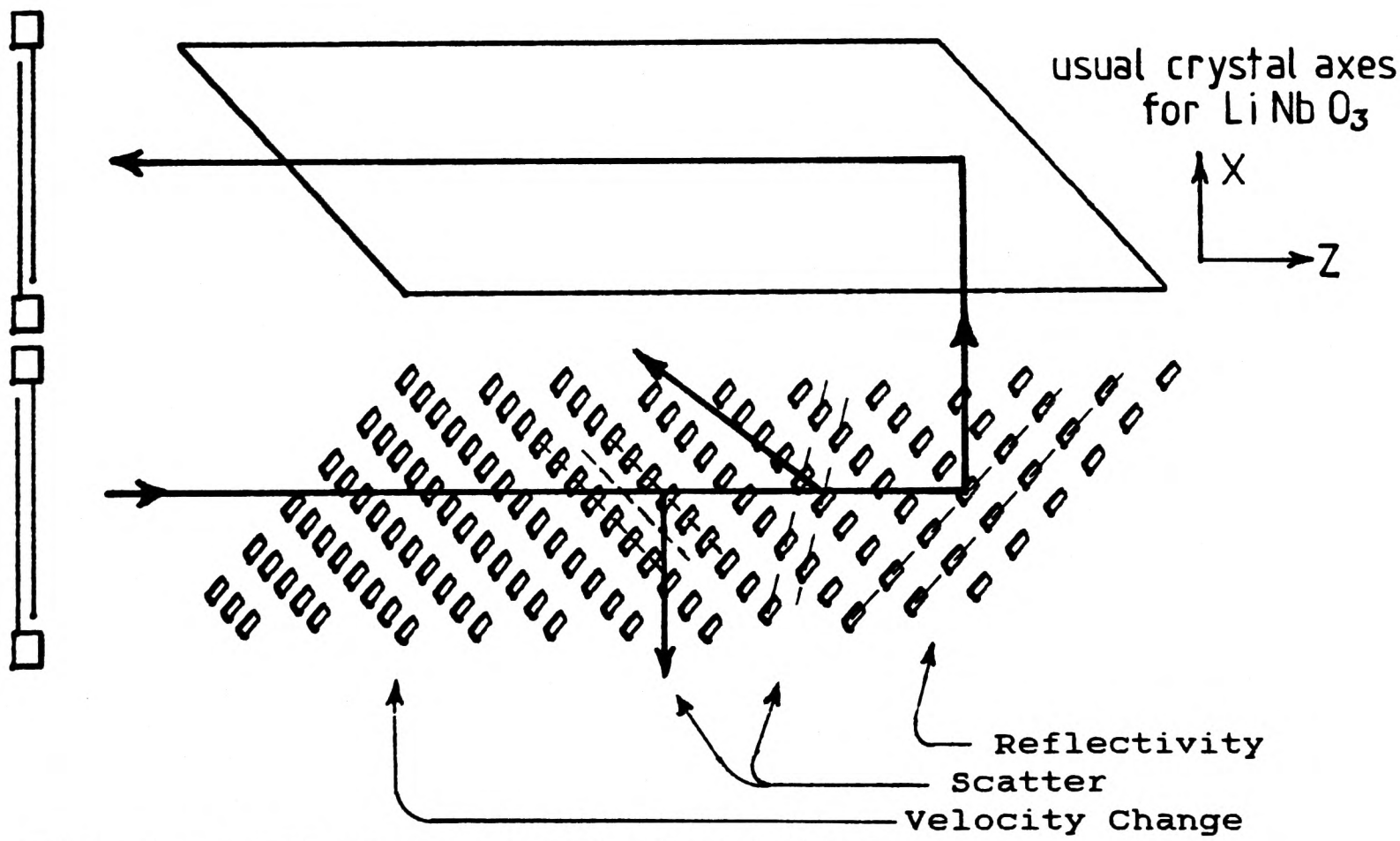


Fig. 1.3: Dot parameters to be studied.

then 'mass loading' is a second reflection mechanism. Finally, the metal dots or strips short out or at least modify the electrical potential at the surface of the substrate. This potential accompanies the SAW on a piezo-electric material, and perturbing it also gives rise to reflections. These three reflection mechanisms are described by Williamson [45,47].

In order to reduce time-sidelobes in the compressed pulse, it is usual to 'weight' the device such that its amplitude response is not uniform with frequency. The amplitude of the output is determined only by those reflectors in the array which reflect synchronously at the frequency in question. An almost arbitary response can be obtained by varying the reflectivity along the array.

The array is usually a few thousand wavelengths long, and the reflectors must be placed to an accuracy of a fraction of a wavelength. In order to do this, the velocity of the wave inside the reflecting array must be precisely known. Velocities also have to be found to determine the angle of inclination of the reflectors to the incident beam. For a reflection through 90°, the inclination is not 45°, but depends on the velocity of the incident and reflected waves, which are not equal on an anisotropic material.

The transducers consist of rows of parallel strips of metal film. Alternate strips are connected to a signal source; the others are earthed. The signal on the transducer applies an undulating electric potential on the surface of the piezo-electric substrate; SAWs are generated if the local

period of part of the transducer is approximately a SAW wavelength. The receiving transducer couples the SAW potential to the external circuit.

1.2 The Aim of this Project

The majority of RACs with high TB products employ depth-weighted grooves as reflectors [11,45,46,47]. The transducers are made of metal film, but the grooves are produced in an ion-beam miller. Two separate processes are required, and the latter one is slow and expensive.

Dot and strip arrays in devices are produced in the same process as the transducers, both being made of metal-film. A mask, with both transducers and arrays, contains all the "information" necessary to produce the device in a one-step photolithographic process. Both strip and dot arrays have been used as a cheap alternative to grooves.

Several weighting schemes are available with strips. The variation of reflectivity with strip width has been studied empirically [19] and used in devices [12,18]. The weighting can also be achieved by strip position: the reflections from a pair of strips can be made to interfere constructively or destructively according to the weighting required. This may also be done with groups of strips [35]. A third method is to vary the length of the strip [22], so that it intercepts only part of the incident wave. Dummy reflectors, which do not reflect in the required direction, are used to fill in the "unused" portions of the array. Most of the devices reported have had low TB products. On a strongly piezo-elctric

substrate, required for high-bandwidth transducers, strips have an inconveniently large reflection. It is difficult to design devices which have effectively a large number of strips reflecting at a given frequency.

This limitation is removed by breaking the strip into a row of dots. Successful dot array devices, including RACs, have been reported. Some were unweighted [14,25]. In others, mass loading was the main reflection mechanism, and reflectivity was either calculated using a simple relation [34,43] or determined by the number of dots per row [36-39,44]. In the only case [48] where dots reflected by electrical shorting, weighting was by the number of dots per row, so that an understanding of the reflections was not required.

It was demonstrated by Stove [40] that the behaviour of a metal-film dot varies radically with its size when it is smaller than one wavelength, if the dots are thin enough for electrical shorting to be the primary reflection mechanism. This is because the potential of small dots floats with the potential of the wave, whereas the potential of a dot longer than a wavelength remains approximately at zero. There is potential for great flexibility when designing RACs and other devices, which has not been exploited previously. The purpose of this project is to study the behaviour of a wide range of sizes and shapes of thin metal dots. It will then be possible to select the optimum size of dot for each particular purpose, rather than to use any arbitrary size, as is done at present. The only other work on this subject is an empirical

measurement of reflectivity [17]; only circular dots are dealt with and even this was not available when the project began. The change in velocity due to dots, which must be known for accurate positioning of the reflectors, has also not been studied before.

Apart from the greater range of dot behaviour, there are two further possible advantages over mass-loading dots. Reflectivity by electrical shorting is relatively independent of metal thickness, reducing the possibility of error. Reflectivity can also be made stronger, so that relatively few dots need to be used. This is useful if an 'electromask' machine is used to generate the mask, as production time and hence cost is approximately proportional to the number of dots.

The theory of dot behaviour will use several concepts originally developed for the design of transducers. A detailed model for tranducer behaviour, in which electrode interactions and the generation of bulk waves are taken into account, was presented by Milsom [27,28]. A simplified version was developed by Morgan [30], who gives relations describing transducer behaviour under the simplifying assumption that bulk waves were not generated. One important difference with dots is that they are not electrically connected as are the electrodes of a transducer. Interactions are either very localised capacitance between near-neighbours or by multiple reflections of the SAW (i.e. interactions of a dot with SAWs reflected by other dots). Multiple reflections can be taken into account with array theory

already being used for grooves and strips [4,31,49]. What is lacking is a knowledge of dots as reflecting elements. The array theory needs to be extended in only one respect. Arrays of dots form lines of reflectors (fig. 1.3) in several directions, equivalent to the several Bragg planes in a crystal. It therefore is necessary to ensure that waves are not reflected strongly in other directions. The problem is aggravated by the fact that mode conversion to bulk waves and leaky surface waves (LSWs) is possible, increasing the number of possible reflections.

It is apparent that three properties of dots have to be investigated (fig. 1.3):

1. The reflectivity of dots, to provide amplitude weighting.

2. Scatter by dots in various directions. The dots do not behave as omnidirectional, isotropic reflectors.

3. Change of velocity in the dot arrays, so that the dots can be accurately positioned. Since the substrate is anisotropic, this has to be done separately for all directions of interest.

In this project, these properties are studied experimentally, although data for scatter from dots is limited. In all cases, the dots are thin enough for the electrical effect to be the dominant mode of reflection. A model is proposed to explain the data, including the effect of capacitance between dots. The reflectivity of dots is

compared with those of holes in a metal film, a new alternative to dots. The model is used on metal strips for comparison with previously-published data. Finally, the results are tested by using them to design a RAC. However, it is hoped that the results will be useful for designing devices in general.

CHAPTER 2

EXPERIMENTAL DETERMINATION OF DOT PROPERTIES

Experiments were conducted to measure the magnitude of reflections of SAWs by thin metal dots on Y-cut $LiNbO_3$, and the velocity change due to the dots. This particular substrate was chosen because it is commonly used in devices, being strongly piezo-electric. However, it is believed that the results could be adapted, with a minimal number of additional experiments, to other materials whose piezo-electric effect is strong enough for electrical shorting to be the main reflection mechanism.

In many devices, such as the RAC (fig. 1.1), the propagation directions are the **Z** and **X** crystal axes, with a smaller portion of the path along **X**. The required reflection is through 90° from **Z** to **X** and **X** to **Z**. The experiments on reflectivity cover a large range of dot shapes. The magnitude, but not the phase of the reflection, was found. In most cases the dots were in sparse arrays, in which interactions were later calculated from the theory to be negligible. A few closer packed arrays were also studied to test the theory on interactions, assumed to be due to capacitatative coupling between dots only.

Also required is a knowledge of the velocity change in the **Z** and **X** directions. When the experiments on velocity were begun, the theory had already been developed, so it was only necessary to confirm it for a small number of dot sizes. Besides, a few potentially useful sizes of dots had been identified from the reflectivity data, and it was thought sufficient to use these. The velocity change in the **Z** direction was verified over a range of frequencies. This is required because a portion of the array such as **P** in fig. 1.1 transmits a band of frequencies, while reflecting at only one. The velocity change in the **X** direction is smaller because of weaker piezo-electric coupling, and in addition the path along **X** is shorter than **Z** in most devices, so velocity change was found only at centre frequency.

Scatter by dots was less extensively studied. Reflections through ±90° are described above. Reflections through 0°, or forward scatter, are deduced from the velocity change. Additional experiments were carried out for 180° reflections (both **Z-Z** and **X-X**). Measured data on scatter is thus limited to these four directions, 0°, 90°, -90° and 180°.

All the "devices" described in this chapter were designed specifically for these experiments; they did not perform any practical signal processing function. The shapes of dots and transducers were defined using the computer language GAELIC (Graphic-Assisted Engineering Layout of Integrated Circuits), and masks made at the Rutherford Laboratories' Electron Beam Lithography Facility using GAELIC software. The devices were prepared from the masks either using the Oxford

SAW group's own facilities or at Plessey Research Laboratories, Caswell.

The theory, and results of both theory and experiment, are presented in the next two chapters.

2.1 Reflectivity (Z to X) of Widely Separated Dots

2.1.1 Layout of devices

Reflectivity of dots through 90° from the **Z** to the **X** direction was measured from periodic, rectangular arrays, each of which had dots of one size. The dots were widely-separated so that interactions would be small. Three similar masks were designed with the general layout of fig. 2.2 . The dimensions are specified in Table 2.1 and fig. 2.1(i). Each mask had six arrays, with different values of dot "length" **a** and dot "width" **b**. (Dot dimensions are specified by their projections on the horizontal axis, as in fig. 2.1.) In most cases, two devices with aluminium thickness of about 200nm were made from each mask to test repeatablity of device fabrication. In addition, devices with aluminium thicknesses of 100nm and 400nm were made from one mask to confirm that mass-loading and topographical effects were small. Except where specified otherwise, **a** and **b** in all the devices described in this chapter were accurate to about 0.5 micron.

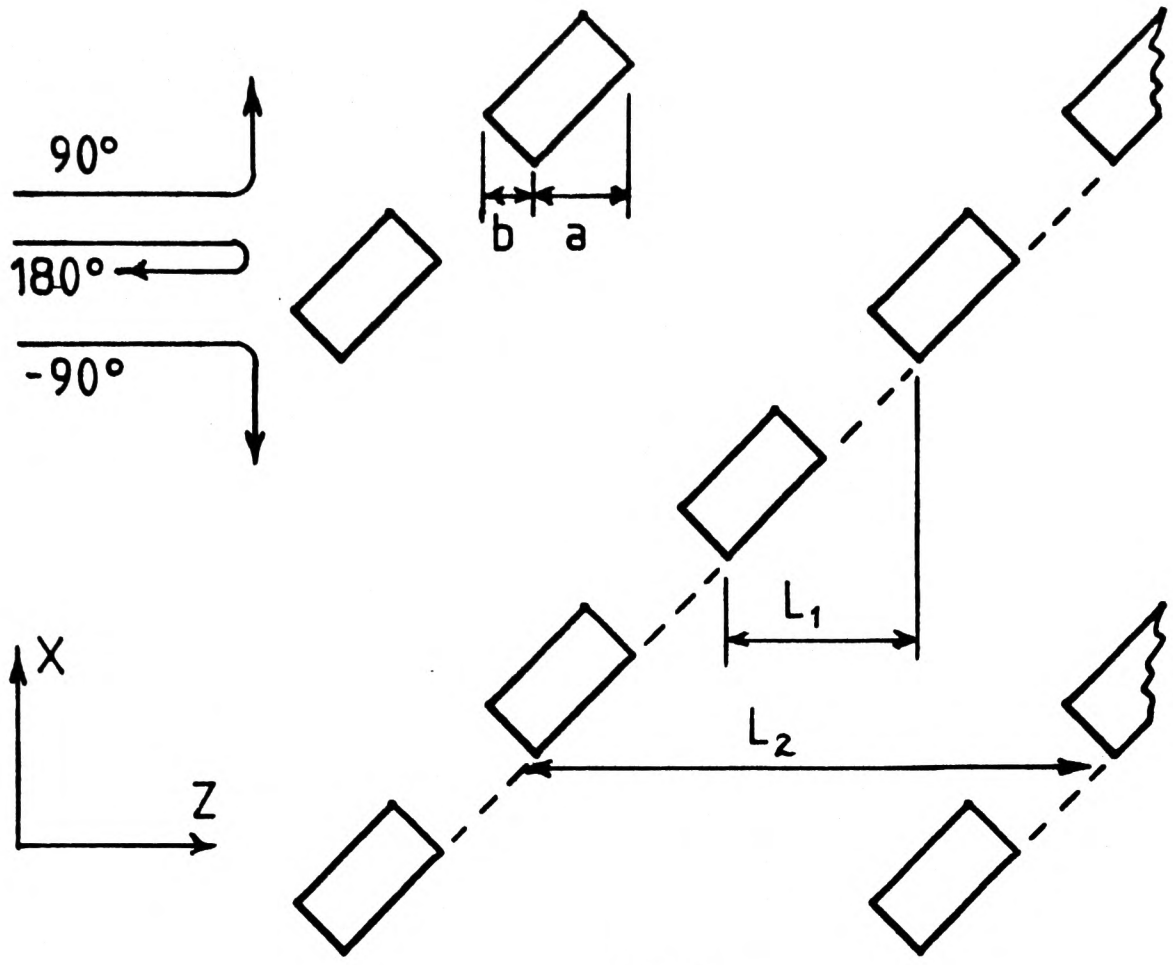
In 5 arrays in each device, the distance L_1 between dots in a row was made equal to 2**a**. L_1 was made different in the sixth array to verify that interaction between dots was small.

Table 2.1: Specifications of Experimental Devices

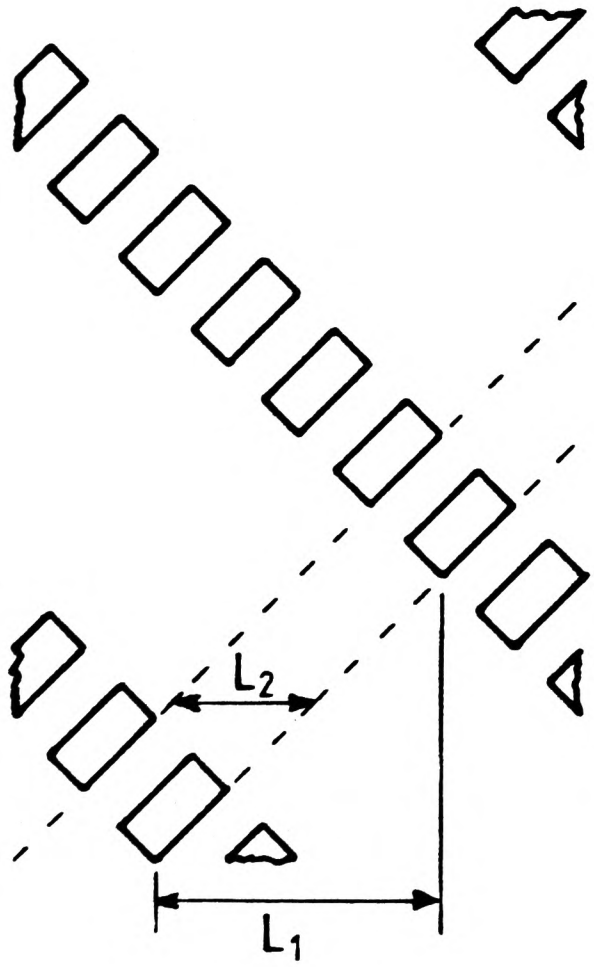
Experiment	Reflectivty of Widely-separated dots: 90, -90, 180 deg.	Accurate Measurement of reflectivity: 90 deg. Velocity Change (X)	Velocity change (Z)	Reflectivity and Velocity change(Z) of additional dot size	Reflectivity of HOLEs, 90, -90 deg
Diagrams: fig...	2(i) 2.2, 2.5	2(ii),(iii) 2.6	2.1(ii),(iii) 2.7	2.1(ii) 2.12	2.1(i) 2.2
DOT SIZE/PATTERN (in wavelengths)			(wavelengths at centre frequency)		
a	0.1 to 1.33	0.5, 1.0, 0.9	1.0, 0.9, 1.0, 1.0	0.25	0.1 to 1.2
b	0.1 to 0.4	0.2, 0.2, 0.3	0.2, 0.3, 0.2, 0.2	0.25	0.1 to 0.6
L1	mostly = 2a	5	2.5 5.0	0.5	= 2.5a
L2	6 or 8	1	1.0 1.0	1.0	4 or 6
"staggering"**	Not relevant	Both	Not S. Both	Not S.	Not relevant
ARRAY SIZE (wavelengths)					
Z	72 or 96	60	200 300	500, 116.5	60 or 90
X	120 or 160	160	30* 40*	50	100 or 150
WAVELENGTH (μm)					
Z	58 or 43.5	70	70 70	58.1	68.2 or 45.4
X	62 or 46.5	75	75 75	--	73.8 or 49.2
TRANSDUCERS					
Centre Freq(MHz)	60 or 80	50	50 50	60	50 or 75
Finger pairs	10	16	4 4	7	10
Tuning coil	No	No	Yes No	No	No
METAL THICKNESS (nm)	100-400	240	240 240	200	200

* Portion of array illuminated by transducers

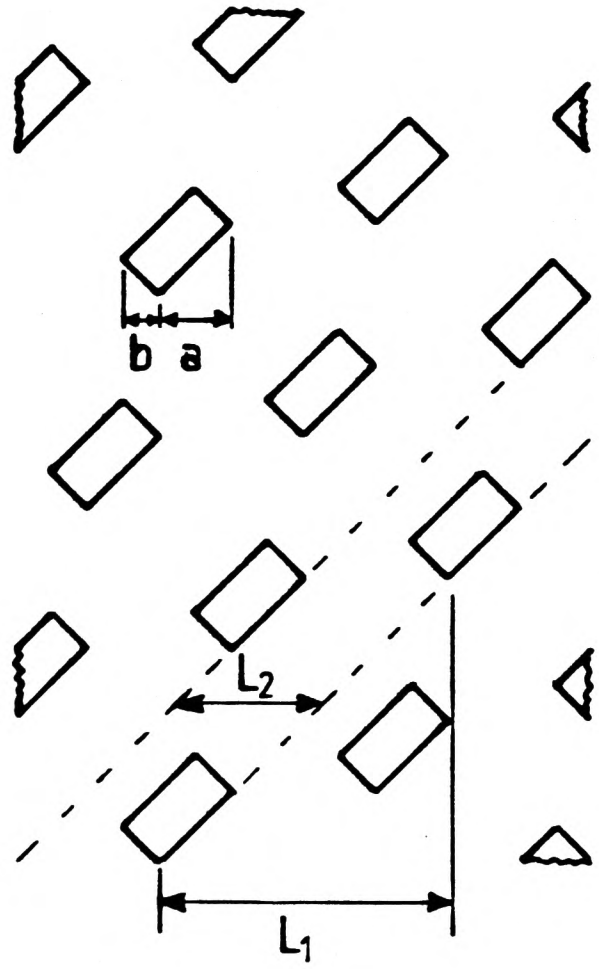
** "Staggered" array: fig. 2.(iii)
"Not staggered" array: fig. 2.(ii)
"Both": fig. 2.(ii) and (iii)
"Not relevant": fig. 2.(i)



2(i)



2(ii)



2(iii)

Fig. 2.1: Dot dimensions and array patterns used. All distances are measured by their horizontal projections.

The wavelengths of the SAW in the **Z** and **X** directions are not equal because the material is anisotropic. Since the arrays were very sparse, free-surface velocities, obtained from reference [1], were used to calculate the wavelengths. The values were rounded to the nearest micron and taken to be 58 and 62 microns respectively at 60 MHz; they are correct to 0.2 microns. The correction required for this error is described later. The rows were thus inclined at 46.9° to the **Z** axis. The sides of the dot, however, were inclined at 45° (fig2.3); this is a limitation of the mask fabrication unit.

The arrays were selected by connecting the appropriate transducers. To find the reflectivity of dots in array A, for example, insertion loss for the four paths shown in fig. 2.2 were found. The straight paths give the insertion loss due to the transducers alone, since array A is weakly reflecting. Arrays B and C do not affect the calculation, being also weakly reflecting and, in addition, affecting paths 2 and 3 equally.

Transducers were placed in pairs around the arrays. Each pair consisted of tranducers with centre frequency at 60 and 80 MHz respectively. Either one was connected to the external circuit. At these frequencies, the rows of dots were 6 and 8 wavelengths apart respectively, so that there was constructive interference between rows. It was therefore possible to measure reflectivity for two values of $(a/\lambda, b/\lambda)$ making 36 sizes of dots from these three devices.

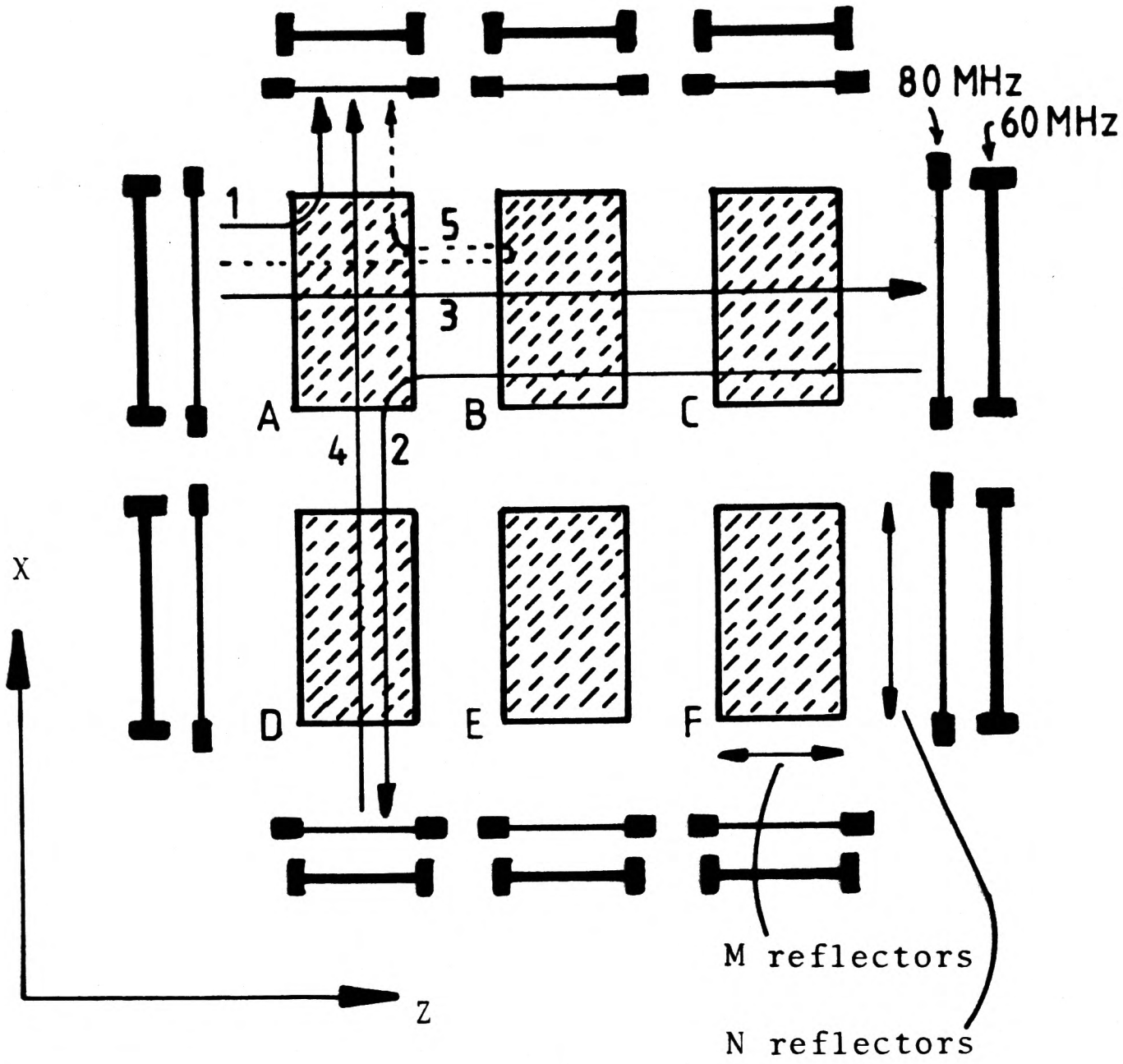


Fig. 2.2: Layout of devices to measure reflectivity.

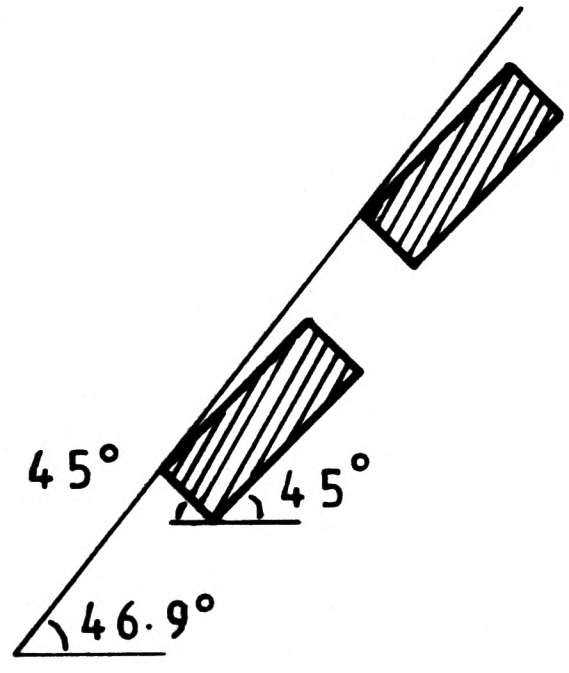


Fig. 2.3: Angle of inclination of dots.

Apertures of the transducers were equal to the array dimensions. The 60 Mhz **Z**-propagating transducer is 120 wavelengths wide, which is approximately the correct aperture for 50 ohms impedance if tuned. The aperture of the **X** transducers was 72 wavelengths at 60 MHz, which is a compromise between a convenient load impedance to the external circuit and the requirement of low diffraction. (The velocity anisotropy near the **Z** direction is such that diffraction is insignificant. This can be verified from data in reference [1].)

The transducers each had 10 finger-pairs, giving a bandwidth of 10%. In a previous, pilot device (which had only two arrays), 4 finger-pair transducers centred at 55 and 75 MHz were used, enabling measurement at 50, 60, 70 and 80 MHz. Bulk waves or leaky surface waves were generated by the **X**-propagating transducers, and these had to be resolved by time gating. This is the reason for reducing the bandwidth of the transducers, at the expense of having fewer measurements. The centre frequencies for leaky and bulk wave generation are higher than for the SAW, so that reducing the bandwidth has the effect of reducing the unwanted waves at the frequency of interest. The problem is also present in the **Z**-propagating transducers [27,28], but to a much lesser extent . The 10 finger-pair transducers operated satisfactorily without a series tuning inductance (to cancel the capacitative part of the impedance), giving an insetion loss of 4 - 9 dB per transducer.

The edges of the substrate were coated with rubber solution or stencil correcting fluid, both found to be convenient SAW absorbers, as a precaution against reflections from the crystal edges.

The same devices were also used for -90° reflections, using the same procedure. This was possible only for some of the arrays, for which constructive interference between dots occured. Measurement of 180° reflections, again using the same devices, is described later.

2.1.2 Measurements

When this project began, an automated system to measure amplitude frequency response had already been developed [40]. It is described in appendix A. The r.f. input is pulsed so that different reflection paths can be resolved by time gating. A 50-point frequency response graph takes only 5 minutes to produce. It was useful to obtain frequency responses for all the paths in fig. 2.2 to check the device, to make the correction described in the next sub-section, and to allow for spurious signals such as path 5 in fig. 2.2, which could not be completely removed by time-gating. These spurious signals, having a longer path length than the main signal, reinforce or subtract from the main signal, causing a ripple in the frequency response. Their effect is removed by averaging. Insertion loss for the L-paths ranged from about 15 dB to 55 dB for the arrays only.

2.1.3 Calculations

Weak reflections are assumed. The reflectivity of symmetric dots is always imaginary [31], and will be denoted jr. If **M** and **N** are the number of rows of dots intersecting the sides of the array as shown in fig. 2.2, then the amplitude of the reflected wave in path 1 is **rN** times the incident wave. Taking the different apertures into account, the fraction of power reflected is $(rN)^2 M/N$. If P_1, ...P_4 are the fractional powers transmitted in paths 1 - 4, then the contribution from the transducers cancel in the relation

$$(rN)^2 \frac{M}{N} = \sqrt{\frac{P_1 P_2}{P_3 P_4}}$$

or

$$r = \left(\frac{P_1 P_2}{P_3 P_4}\right)^{1/4} \frac{1}{\sqrt{MN}} \qquad (2.1)$$

The relation is therefore insensitive to minor faults in the transducer, an important feature of this method of measurement. It should be noted that equation (2.1) implies that amplitudes of the SAW are defined so as to be proportional to the square-root of the power density. The relation between amplitude so defined and elctric potential or particle velocity depends on the properties of the anisotropic material in the particular direction.

Because the assumed values of wavelengths were rounded to the nearest micron, the angle of inclination of the rows to the **Z** axis, 46.9° (fig. 2.3), was slightly inaccurate and a correction was made for this. If f_z is the frequency at which

all the reflectors on a line parallel to the **Z** axis are one wavelength apart, and f_x is similarly defined, then, as derived in Appendix E,

$$\sqrt{\frac{P_{out}}{P_{in}}} = r\sqrt{MN}\,\frac{\sin(\pi M f/f_z)}{M\sin(\pi f/f_z)}\,\frac{\sin(\pi N f/f_x)}{N\sin(\pi f/f_x)} \qquad (2.2)$$

Because of the slight error in the angle, $f_z \neq f_x$, and the maximum value of the above expression when **f** is varied is less than $r\sqrt{(MN)}$. f_x and f_z were estimated by fitting equation (2.2) to the frequency response of the device (paths 1 or 2). An example of the device response appears in fig. 2.4. In all cases described in this chapter, the correction was less than 7%.

2.2 180° Reflections by Dots

Although the devices described in the previous section were not designed for 180° reflections, satisfactory results were obtained from them. Both **Z-Z** and **X-X** reflections were measured for those arrays for which the condition of constructive interference between dots was satisfied.

Two procedures were used (fig. 2.5). In the first, the insertion loss of path 7 (or similar) was measured, while that of path 6 was already available from previous measurements. Since the reflectivity of array F through ±90° was already known, 180° reflections from array E could be found. Because of the finite impulse response of array 6, path 7 could not be completely resolved from path 8 by time gating; this is one source of error.

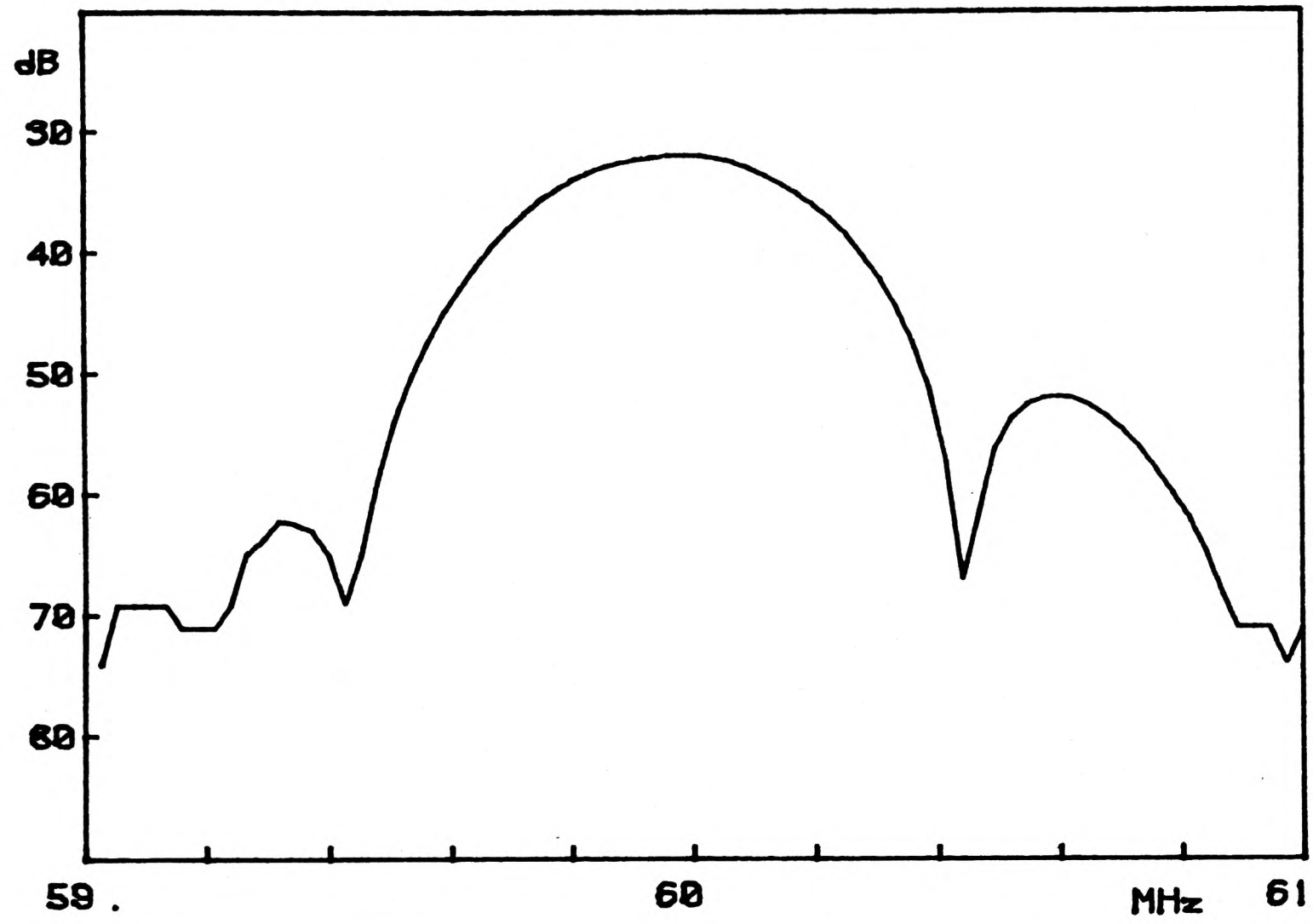


Fig. 2.4: Insertion loss of the device in fig. 2.2 in L-path, i.e. a path similar to no.1

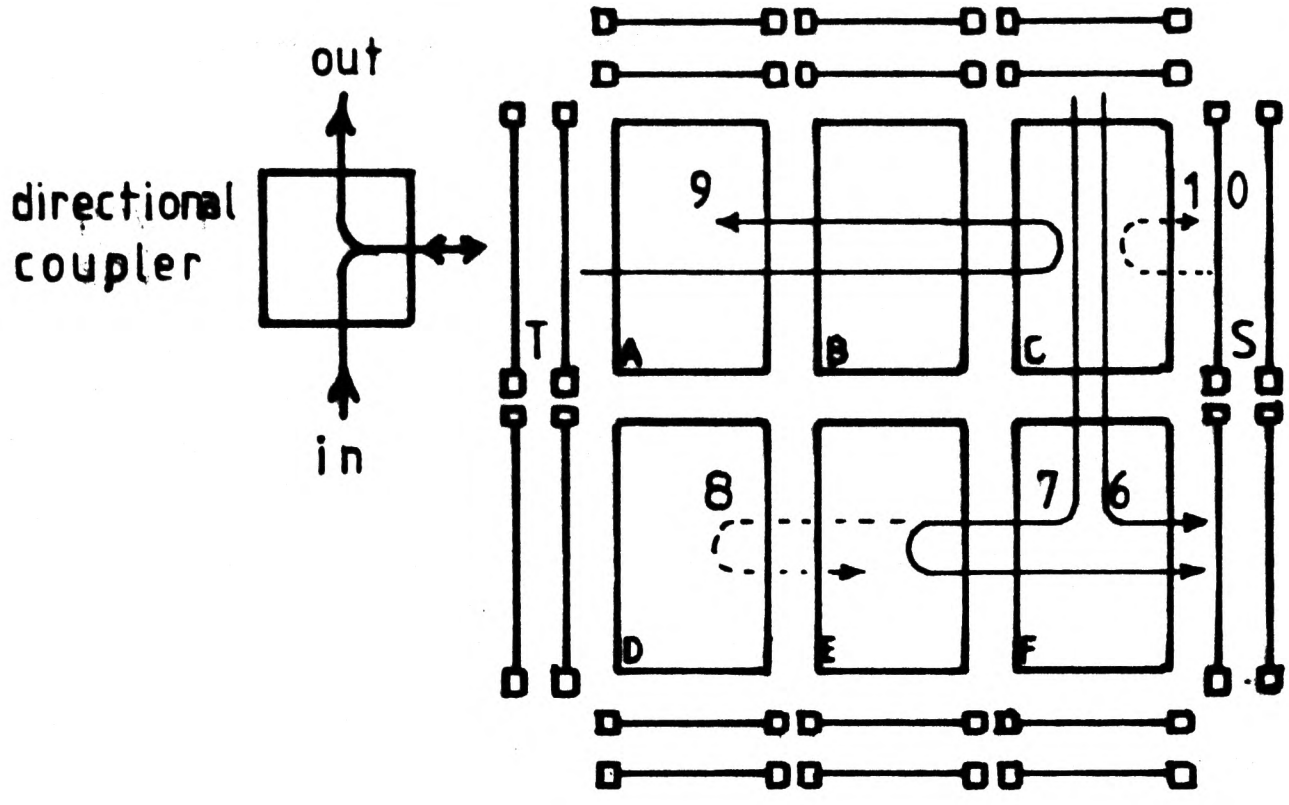


Fig. 2.5: SAW paths for 180 deg. reflections.

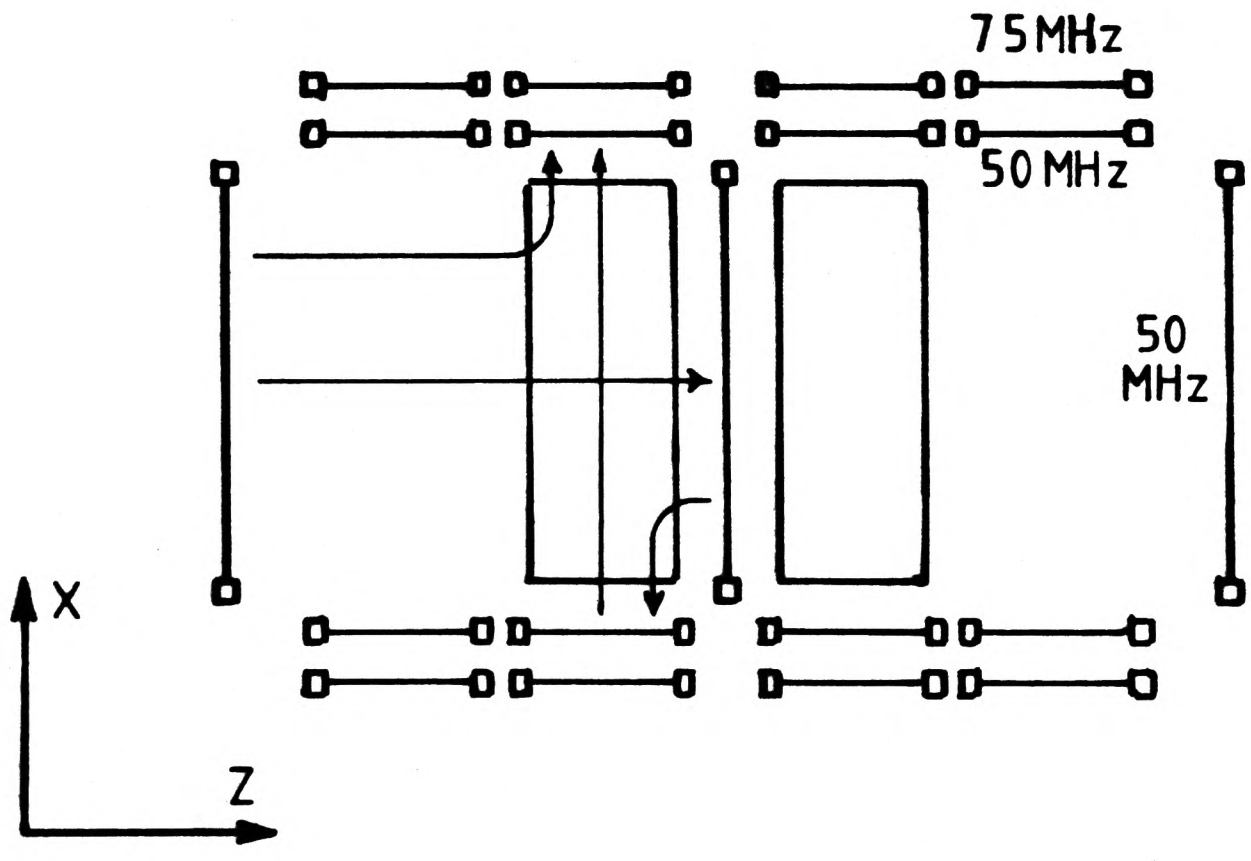


Fig. 2.6: Devices for measuring reflectivity of dots with near-neighbour interactions, and measuring X-velocity change

The other method empolyed a Hewlett Packard 8502A transmission/reflection test set as an external directional coupler to separate input and output signals in paths such as no. 9. The response of transducer T was found from path T-S, assuming that transducers T and S were identical. Measurement of path 10 was not attempted because the reflected wave arrived back at the transducer a very short time after the end of the transmitted pulse; it was feared that the system's amplifiers would still be saturated by leakage of the transmitted pulse in the directional coupler.

The calculations required are similar to those for 90° reflections, except that **rM** (for **Z-Z** reflection) or **rN** (**X-X**) replaced **r√(MN)** . The results from the two procedures do not differ significantly in that the difference was within the scatter of points obtained from one procedure. Errors in the angle between the row and the propagation direction are irrelevant to this measurement.

2.3 Further Experiments on Reflectivity (Z-X), Including Dots with Near Neighbours

When a few potentially useful sizes of dots had been identified, it was decided to measure their reflectivities more accurately. For each dot size, measurements for two types of arrays were made (fig. 2.1(ii),(iii)), so that the theory for interactions between dots could also be verified. The theory (chapter 3) predicts negligible interactions for the pattern of fig. 2.1(iii), but significant interactions for some dot sizes in the pattern of fig. 2.1(ii). It should be noted that, with L_2 small and L_1 large, the dots can have

near-neighbours even though the array is relatively sparse.

The devices used had the layout of fig. 2.6, and further details are listed in table 2.1. The 50 MHz transducers not facing an array and all the 75 MHz transducers are for a different experiment. It was convenient to have two arrays per device. The smaller number of transducers and arrays on the device (compared with those described in section 1 of this chapter) was necessary to reduce the number of spurious reflection paths. In addition to the stencil-correcting fluid used as SAW absorber on the edges of the device, distilled water was used as a temporary absorber over those transducers and the array which were not being used in each measurement. Spurious reflections were kept to a minimum in this way.

The measuring equipment used and the calculations were similar to those described in section 1 of this chapter. Insertion loss per array was between 9 and 20 dB. Multiple reflections were neglected in the calculations resulting in only a small error.

2.4 Reflection by Holes

Structurally, holes are the exact opposite of dots: the substrate is completely covered with metal except in the holes, where it is free. They are an alternative type of reflector to dots, and will become the topic for a separate research project by another worker. Reflectivity by sparse hole arrays only are treated here.

The mask designed is similar to that shown in fig. 2.2, except that the dot arrays are replaced by rectangular panels of metal film with holes. Dimensions appear in table 2.1. The number of dots per row was reduced to limit interactions between holes further. Both the devices made were accidentally over-etched by about 3 microns, which is the reason for the unusual dot sizes presented in the results. This overetching also changes the mark-to-space ratio along the row; this was compensated for by simple scaling. Except for this, the measurements and calculations are the same as in section 1. Reflections through ±90° were measured. From the -90° data, equivalent 90° reflections were calculated by interchanging length **a** and width **b** with aprropriate scaling to account for the number of dots per row. All the data is presented in chapter 4.

2.5 Velocity Change in Dot Arrays: Z direction

2.5.1 Device Layout

Velocity change due to dots was measured from the phase lag of a dot array relative to a reference channel (fig. 2.7). Velocity changes due to temperature and different batches of crystals affect both channels by about the same amount, so they largely cancel. Each crystal substrate was large enough for two such devices. In order to reduce multiple reflections, the arrays were made as narrow as possible, subject to diffraction limitations. Each array was divided into two halves. The only significant reflections in the top half (fig. 2.7) are upward, and the lower half was designed to reflect downwards. (This was determined from the theoretical

model.) The number of possible multiple reflection paths (such as 1 and 2 in fig. 2.7) were further reduced in this way. The arrays were made 20 wavelengths wider than the transducer to reduce diffraction along the edges of the array. The dot sizes and array pattern are listed in fig. 2.1 and table 2.1. At each end of the array and reference channel were placed two 4-finger-pair transducers, both centered at 50 MHz. As in previous devices, the edges of the crystal were covered with stencil-correcting fluid.

2.5.2 Measurements

Phase lag was measured with an HP 8754A Network Analyser. Frequency range was limited only by the bandwidth of the transducers. At first, only the inner transducer at each end was used, the other being covered with absorber. In a later measurement, the absorber was removed and each pair connected in parallel to from a short "ladder" transducer. The result was that signal was increased relative to the direct electromagnetic coupling between input and output. It was then possible to make measurements at frequencies in the lower sidelobe of the transducer, increasing the frequency range available. The higher sidelobe was not useful because of bulk wave generation.

The range of phase shift $\Phi°$ measured ranged from between about 0.5 and 1.5 cycles. It varied only slowly with frequency so not many readings were required; these were quite conveniently taken manually either from the graph produced by an X-Y analogue plotter connected to the Network Analyser, or directly from the CRT display. The measurements

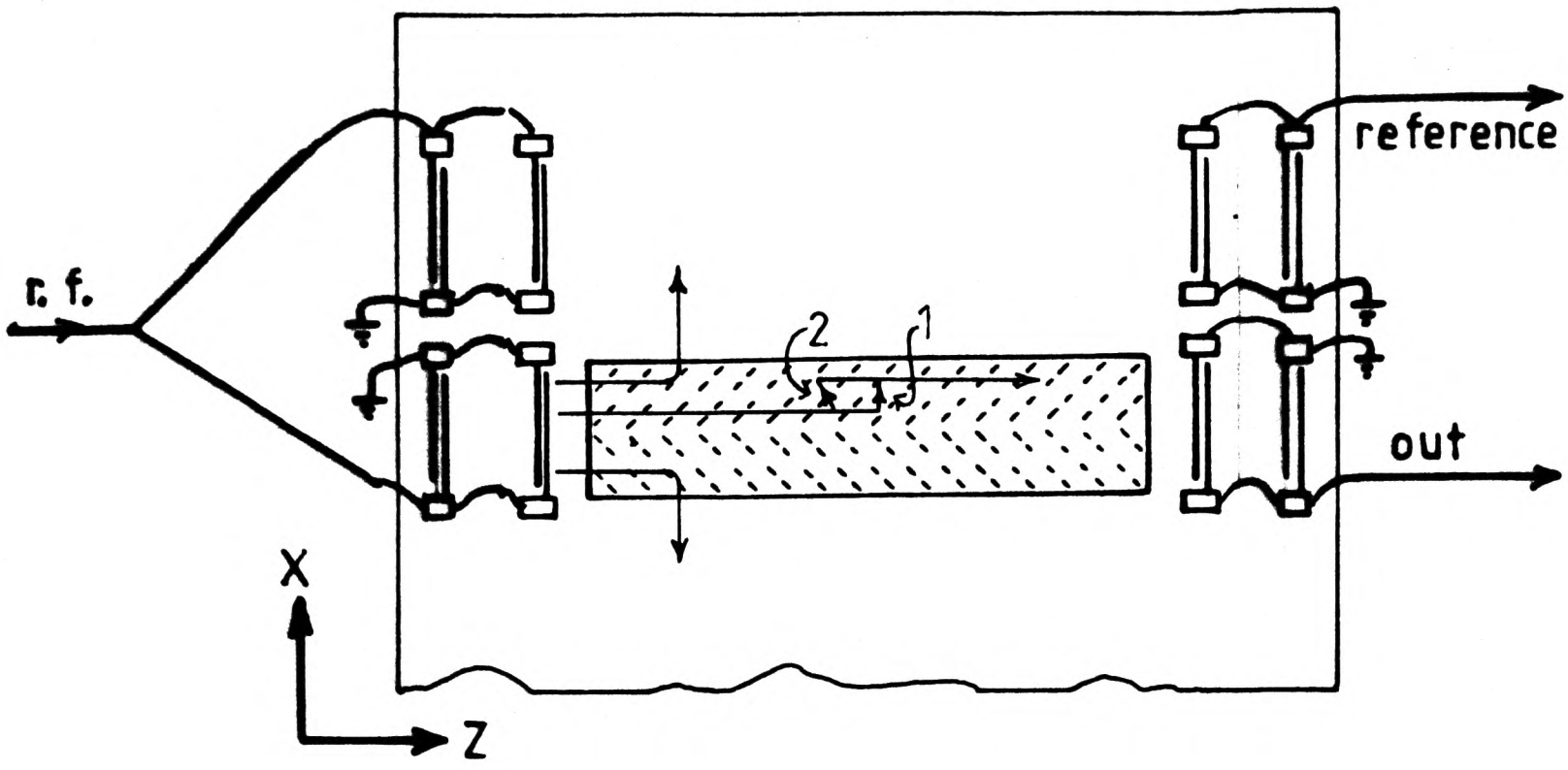


Fig. 2.7: Device for measuring velocity change in the Z-direction. Connections for phase measurement are indicated.

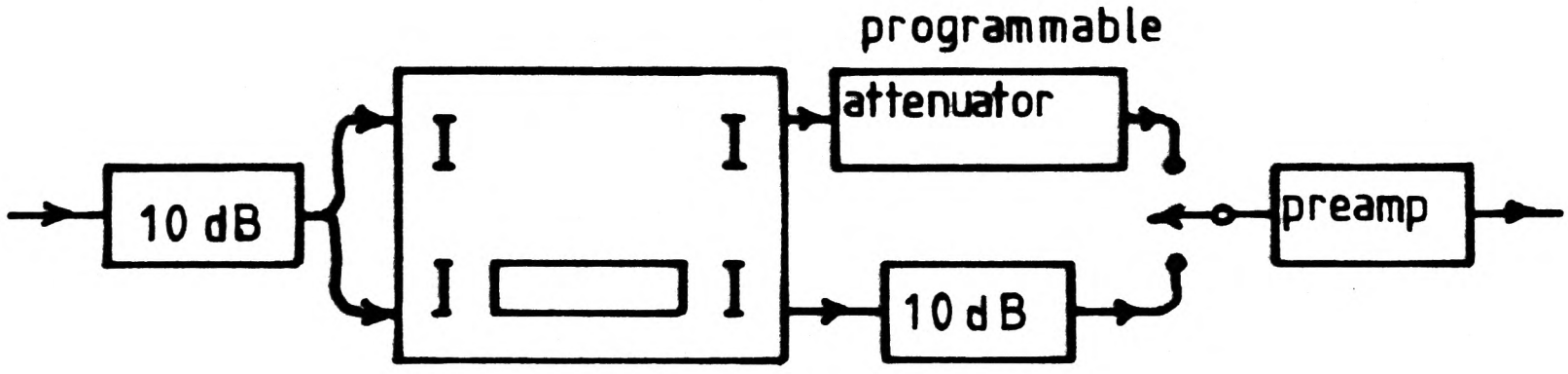


Fig. 2.8: Modifications to the insertion loss circuit (Appendix A) to cancel out transducer response.

cannot distinguish between (say) $\Phi°$ and $(\Phi \pm 360)°$. However, for the assumed values $\Phi°$, $(\Phi - 360)°$ would have been negative (phase lead rather than lag) for part of the frequency range, implying an increase in SAW velocity rather than slowing. $(\Phi + 360)°$ would have been improbably large.

Since there were two reference channels per device, it was possible to measure the phase difference between them, to obtain an estimate of the likely errors due to inconsistencies in the transducers and the substrate. Up to about 8° or 0.02 cycles phase error was measured.

The insertion loss of the array channel relative to the reference channel was also measured to monitor the level of multiple reflections. The insertion loss measurement system (Appendix A) was modified as shown in fig. 2.8 to get a direct reading of the ratio of the insertion losses.

2.5.3 Calculations

It is easily shown that the fractional change in velocity due to dots is

$$\frac{\Delta v_d}{v} = \frac{-\Phi}{360} \quad \frac{f_o}{Mf} \qquad (2.3)$$

where Φ is the phase lag in degrees and Mf/f_o is the number of wavelengths of the array at frequency f.

A typical plot of phase lag is given in fig. 2.9 . Theory predicts that the reduction of SAW velocity is a slowly-varying function of frequency. The very rapid ripple is due to direct electromagnetic coupling and triple transits.

Both these spurious signals are removed by averaging. It was sufficient to estimate the average by eye. The undulations in the phase shift are due to multiple reflections through about 90°. The multiple reflection path is up to about 15 wavelengths longer than the reference path. The difference in impulse lengths enables the frequency response to vary rapidly. (Reflections in other directions were predicted theoretically to be small.) Multiple reflections are evidently the largest source of error despite the design features intended to reduce them. Although the correction turned out to be less than about 10%, it was worthwhile to compensate for their effects. Reflectivity was first estimated from each insertion loss plot, such as that shown in fig. 2.10 . Insertion loss is approximately constant with frequency except for one or more notches, where the scattering condition is satisfied by one of the array scattering vectors. (Scattering conditions are described in chapter 3 .) Reflectivity was estimated from the depth of these notches. The phase error caused by double reflections was then estimated, and subtracted from the phase shift data. This correction was made only in those cases where the insertion loss notch was greater than about 1 dB. The worst notch was about 7 dB deep.

The following derivation of the correction assumes weak reflections, so that only double-reflections are considered. Each row of dots in a device consists of only three or four dots. In order to avoid the question of whether each row behaves exactly like a strip even at short range, it is convenient to consider reflections per "column" of dots

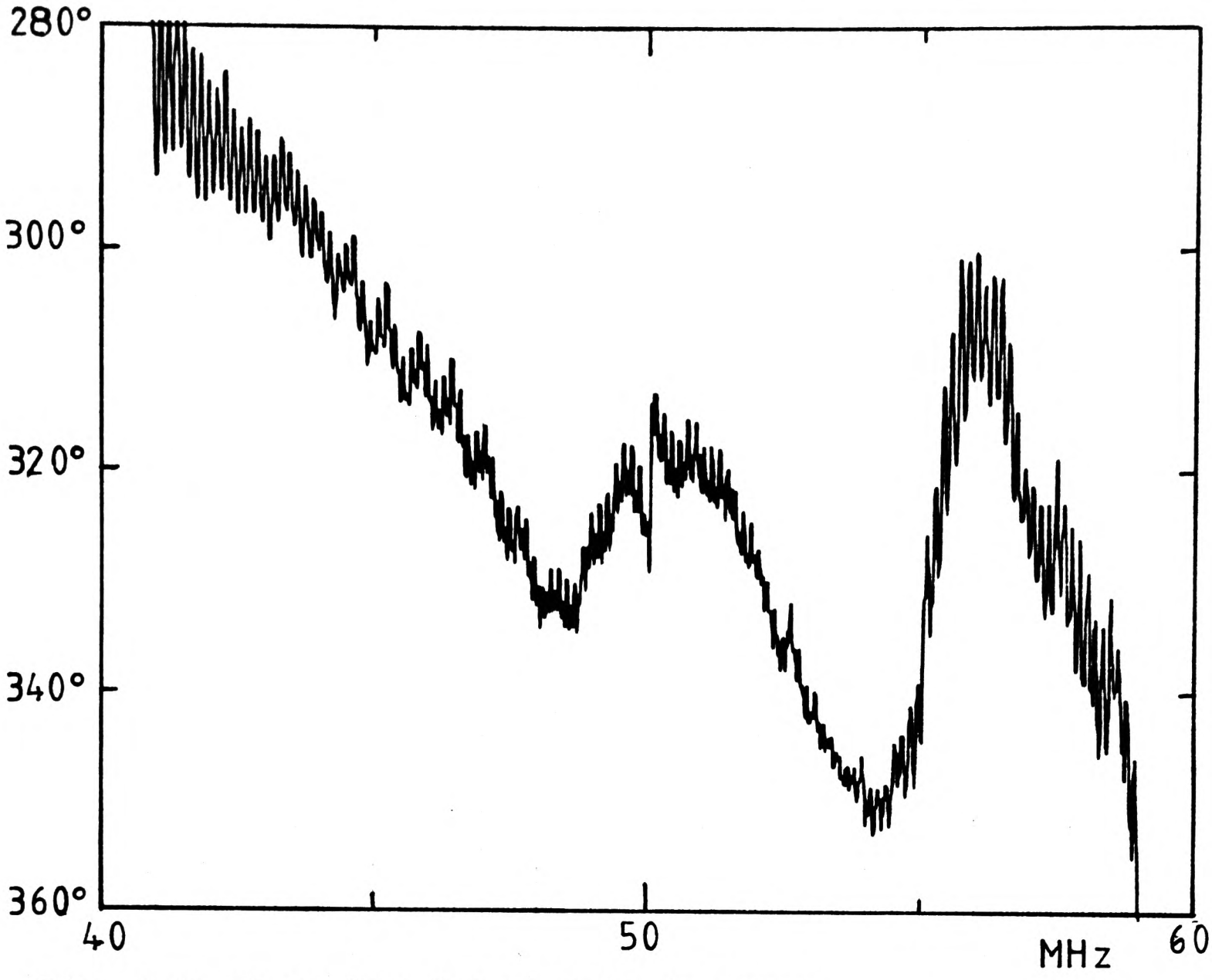


Fig. 2.9: Phase lag due to dots in array

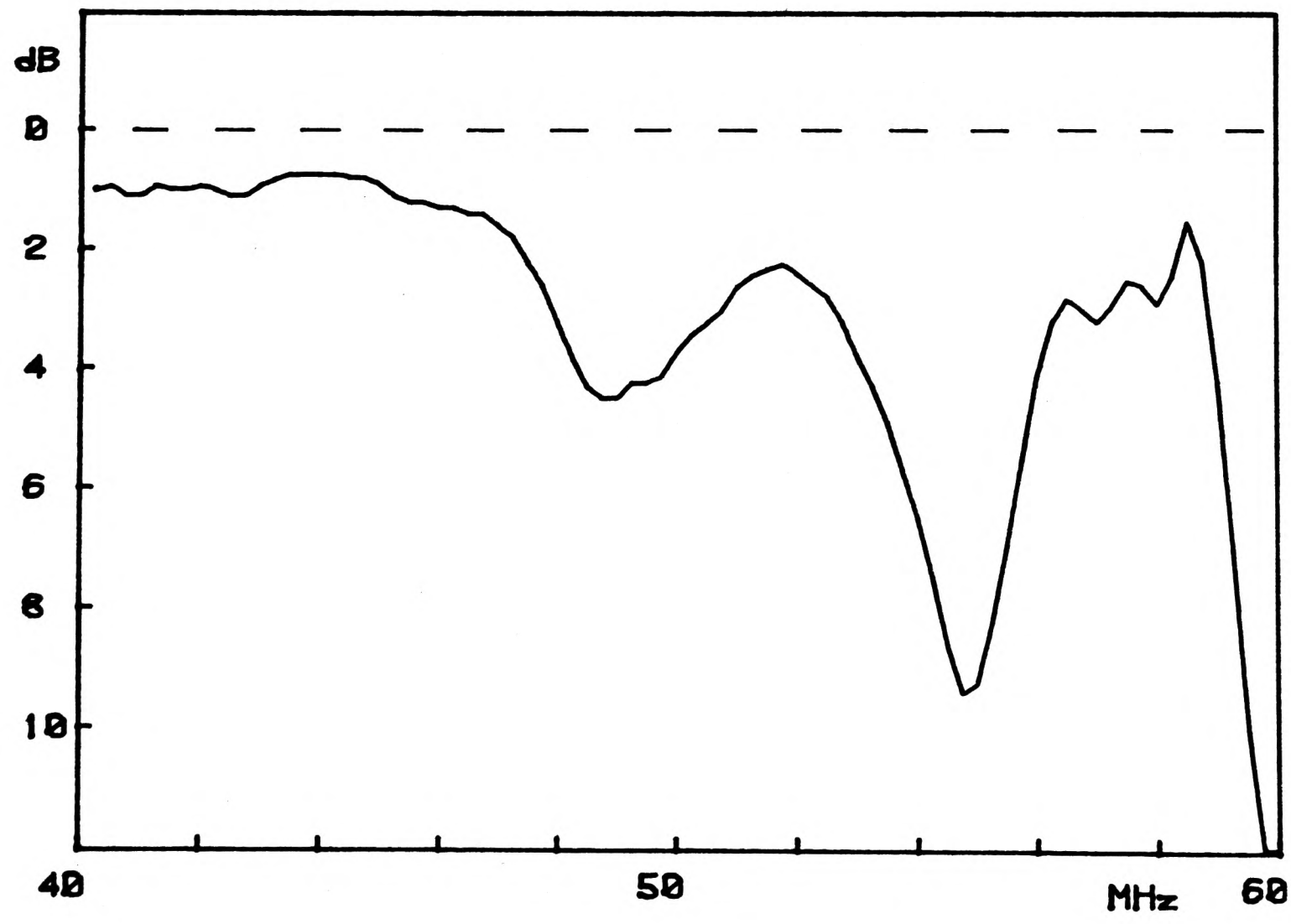


Fig. 2.10: Attenuation due to dots in array

(fig. 2.11). There are 20 to 40 dots per column. The angle ξ through which the beam is reflected is a function of frequency, and is such that there is a path difference of an integer number of wavelengths between dots in the same column.

The error in output signal due to double reflections relative to the reference channel is found by summing all possible reflection paths. Including also the depletion of power $0.5r_{col}^2$ per column,

$$V_{err} = -M_{col} \sum_i r_{col}^2 \frac{W_i}{W} \exp(-j2\pi \frac{l_i}{\lambda}) - \frac{1}{2} M_{col} r_{col}^2 \qquad (2.4)$$

where M_{col} is the total number of columns in the array, and jr_{col} is the reflectivity of a column of dots. The other parameters are defined in fig. 2.11 . W_i and l_i are most conveniently found geometrically on graph paper. r_{col} is found by fitting the notches of $Re(V_{err})$ with the insertion loss plot; $Im(V_{err})$ is then a correction term for the phase plots. Additional remarks on double reflections and depletion of power per column are made in Appendix F.

It should be noted that stop bands are not present in the data. They are caused by multiple reflections which have been minimised in the design and further reduced by the compensation just described.

The value of r_{col} determined above does not agree with theoretical values nor with other experiments, being about 2.5 times too large. It is believed that a reflection from SAW to

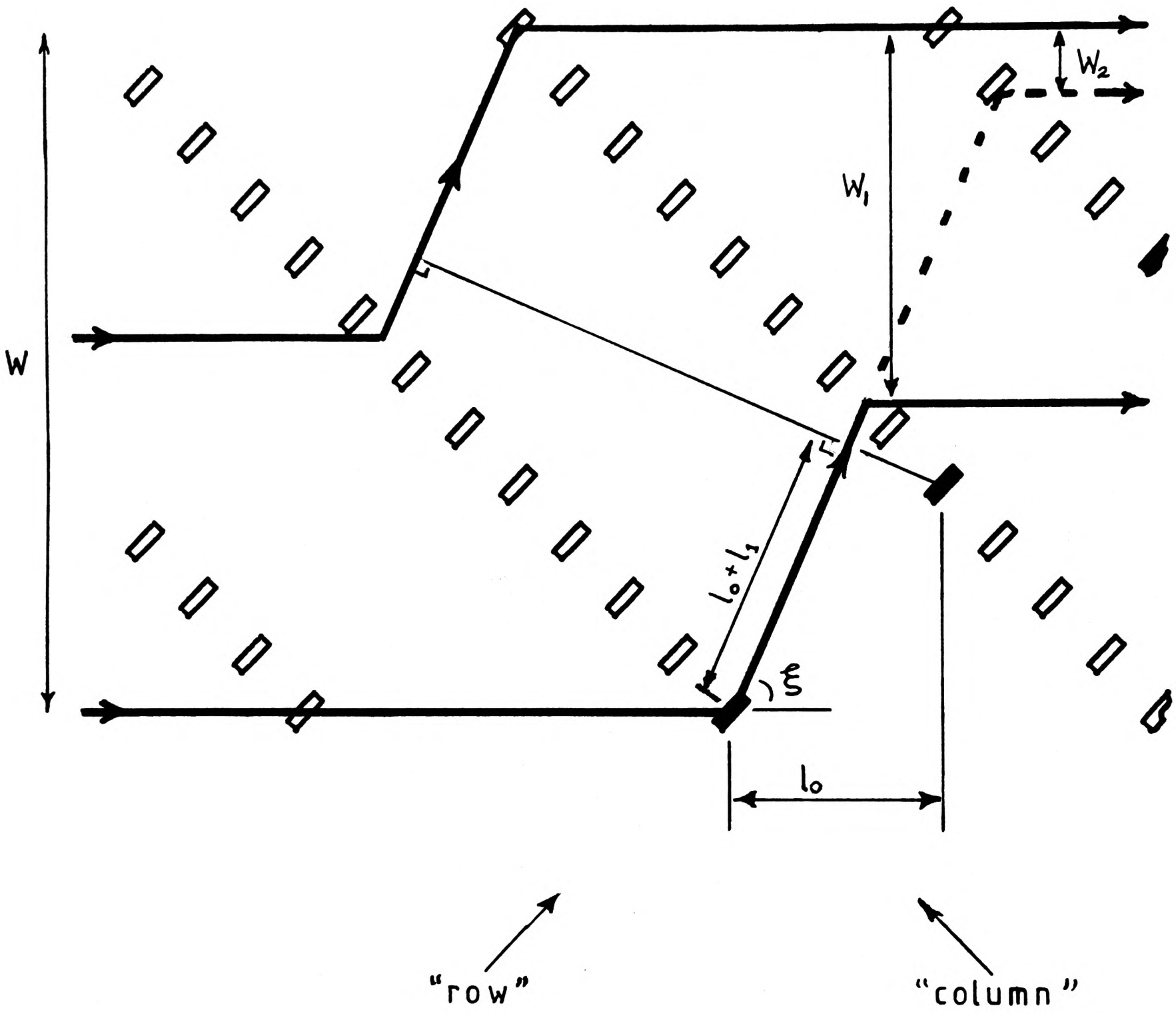


Fig. 2.11: Calculation of phase error due to multiple reflections.

LSW to SAW is involved. As explained in chapter 3, the array has a large number of "scattering vectors", each corresponding to a set of Bragg lines. It is possible that the scattering conditions (chapter 3) for the SAW to LSW conversion are satisfied by one of these scattering vectors at approximately the same frequency as the 90° SAW to SAW reflection, and that the average path lengths of the SAW-LSW-SAW and SAW-SAW-SAW signals are similar. It would then be difficult to distinguish between the two modes of reflection.

2.6 Velocity Change in Dot Arrays: X Direction

The devices described in section 3 of this chapter (fig. 2.6) were also designed for velocity change in the X direction. The **X** propagating transducers have 16 finger-pairs each to reduce leaky wave generation at centre frequency. It was not considered appropriate to fabricate devices from masks designed for velocity change in the **Z** direction because the transducers had wider bandwidth, and also because diffraction is more serious in the **X** direction, necessitating wider arrays. The procedure for measurement and the calculations are the same as in the previous section, except that no significant notches were found in the insertion loss, so compensation for double reflections was not required.

The devices in fig. 2.6 also had 75 MHz **X**-propagating transducers. At this frequency, the value of (a/λ,b/λ) for one of the devices was (1.35,0.45). In addition to the dot sizes quoted in table 2.1, the velocity of the **X** propagating SAW was measured for this larger dot size without the need for further devices.

2.7 Further Experiments on Velocity and 180° Reflections

2.7.1 Device Layout

Devices from an existing mask were found to be suitable for measurement of 180° reflections and velocity. The mask was designed for a different purpose by Stove [40], and made by RSRE Malvern. Being from a different source, it serves as a check for the Rutherford masks, while also yielding data for one additional size of dot.

Details of the device appear in table 2.1 and fig. 2.12 . It had a blank channel (path 1) and two channels with dot arrays 500 wavelengths long (paths 2 and 3). Part of the array in path 3 was etched away, giving the shorter, 116.5 wavelength array illustrated.

2.7.2 Measurements

Because of a breakdown in the insertion loss equipment, both insertion loss and phase lag (as functions of frequency) were found from the network analyser. The insertion loss of paths 2 and 3 were found relative to reference. At resonance, part of the incident wave is reflected by the array in path 3, while at other frequencies, it is almost transparent. The depth of the notch in the amplitude reponse in path 3, 14.7 dB, therefore gives the amplitude transmission coefficient of the array. The array in path 2 was not suitable for this measurement since the transmitted wave was too small to measure, but 180° reflections could be calculated from the width of the stop

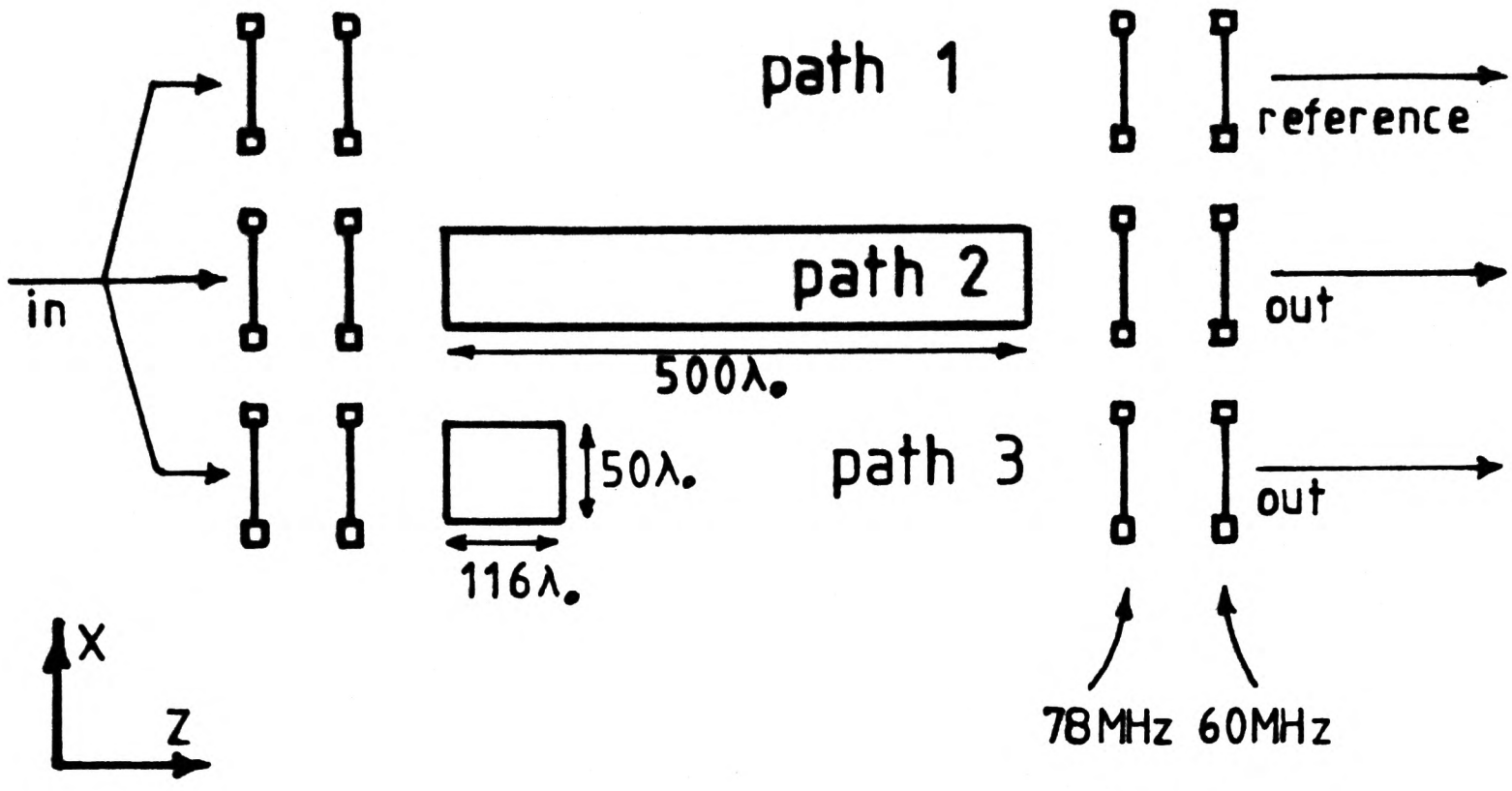


Fig. 2.12: Additional device for velocity change and 180 deg reflections.

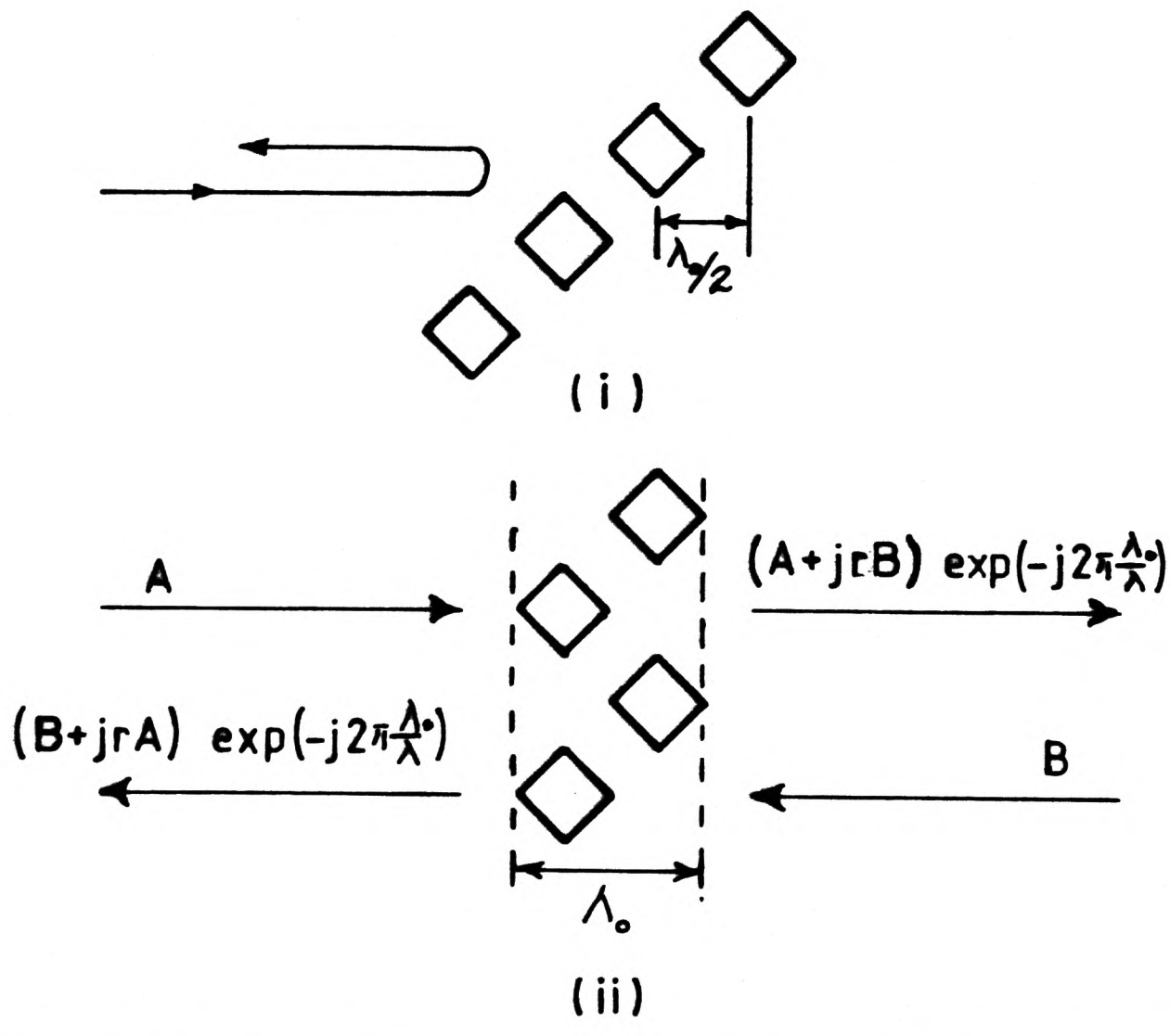


Fig. 2.13: Calculations involved for the device in fig. 2.12.

band. The phase shift of the array in path 2 was measured as in section 5. It ranged between 1.0 and 5.5 cycles. A different correction for multiple reflections was used, since 180° and not 90° reflections dominate (as found both theoretically and experimentally).

A slightly wider bandwidth was available by separately utilising both 60 and 78 MHz transducers. Although the transducers were relatively narrow-band, having 7 finger-pairs each, satisfactory results were obtained at frequencies inside two of the lower sidelobes and one upper sidelobe of the transducers.

2.7.3 Calculations

Reflectivity is calculated from the amplitude transmission coefficient using a relation given by Otto [31, with different notation]. The transmission coefficient is $1/\cosh(\mathbf{M}r)$.

The following derivation was used to find reflectivity from the width of the stop band, and to correct the phase response for multiple reflections. The reflectivity of the row of dots in fig. 2.13(i) at resonance is equal to that of the cell of dots in fig. 2.13(ii). The reflectivity of the cell varies very slowly with frequency. Terms with powers above r^2 are neglected.

In the notation of fig. 2.13, λ is the wavelength at a particular frequency and λ_o is the period of the array. λ is given by $v = f\lambda$, where v is the velocity of the SAW in the array. At frequencies within the stop band, arg[$(A+jrB)$

$\exp(-j2\pi\lambda_o/\lambda)$] is one cycle.

Over one cell,

$$\Delta A = A\left[\exp\left(-j2\pi\frac{\lambda_o}{\lambda}\right) - 1\right] + jrB\exp\left(-j2\pi\frac{\lambda_o}{\lambda}\right) \qquad (2.5)$$

and a similar relation exists for ΔB. Since $\exp(-j2\pi\lambda_o/\lambda) = \exp(-j2\pi(\lambda_o-\lambda)/\lambda)$ and $(\lambda_o-\lambda)/\lambda$ is small near the stop band,

$$\exp\left(-j2\pi\frac{\lambda_o}{\lambda}\right) = 1 + j\phi \qquad (2.6)$$

where

$$\phi = 2\pi\frac{\lambda_o-\lambda}{\lambda} = 2\pi\frac{f/v - f_o/v_o}{f_o/v_o} \qquad (2.6a)$$

v_o is the free surface SAW velocity, and f_o is defined as being equal to v_o/λ_o

Over Δn cells,

$$\left.\begin{aligned}\frac{\Delta A}{\Delta n} &= j(-\phi A + rB)\\ \frac{\Delta B}{\Delta n} &= j(-rA + \phi B)\end{aligned}\right\} \qquad (2.7)$$

If

$$A = A_o\exp(-j\theta n) \qquad (2.8)$$

then

$$\frac{\Delta A}{\Delta n} \simeq \frac{dA}{dn} = -j\theta A \qquad (2.9)$$

and similarly for B.

So

$$\begin{bmatrix}(\theta-\phi) & r\\ -r & (\theta+\phi)\end{bmatrix}\begin{bmatrix}A\\ B\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix} \qquad (2.10)$$

For non-trivial solutions, the determinant for the matrix is

zero, resulting in

$$\phi^2 = \theta^2 + r^2 \qquad (2.11)$$

If ψ is the phase lag of the array channel relative to the reference, which is measured as a function of frequency in the experiment,

$$\theta = 2\pi \frac{f - f_o}{f_o} + \frac{\psi}{M} \qquad (2.12)$$

where the first term corresponds to the phase lag in the reference channel. **M** is the number of cells in the array. (2.6a), (2.11) and (2.12) are sufficient to find **v**.

At the edge of the stop band, $\theta=0$, (Re $\theta=0$, Im $\theta=0$) so using (2.11) and (2.6a), jr is found from the two stop-band edges f_s in the relation

$$\pm r \simeq 2\pi \frac{f_s - f_o}{f_o} \qquad (2.13)$$

2.8 Velocity of the Leaky Surface Wave

The velocities of the Leaky Surface Wave through free and metallised surfaces are required in the theory for SAW reflections described in the next chapter. These values have already been measured by a time-of-flight method [20], and also calculated from the material constants [20]. It was decided to make a more accurate measurment using phase. The device is shown in fig. 2.14 . The path length is short because the LSW on free and metalised surfaces decays by radiation into the bulk at about 0.25 and 0.6 dB per

wavelength respectively [20]. The path difference between the metallised or free channel and the reference channel is 1.8 mm or about 24 wavelenths at 55 MHz, which is the transducer centre frequency for LSW. The transducer aperture was 4.2 mm.

SAW generation is negligible in the pass band for LSW generation. Even so, it was not possible to measure the phase difference between the free or metallised channel and reference using the method described in section 5 because of the presence of bulk waves. Instead, pulsed measurements were made with the circuit described in Appendix B, so that most of the bulk waves could be removed from the measurement by time gating. A pulse consisted of about 50 cycles of r.f.

The output of the phase sensitive detector (mixer) is zero at those frequencies at which the output from the device is $(360n+\pm90)°$ out of phase with the input. These frequencies were found manually in three separate experiments for the free, metallised and reference channels respectively. Phase shift at other frequencies could be found by interpolation. The phase difference between the free or metallised channels and the reference was then calculated.

Velocity was calculated using a method similar to the velocity measurements in dot arrays. As in section 5, it is not possible to distinguish between $\Phi°$ and $(\Phi+360n)°$, so an approximate value of $\Phi°$ was required from another source. This was calculated from the data provided by reference [20].

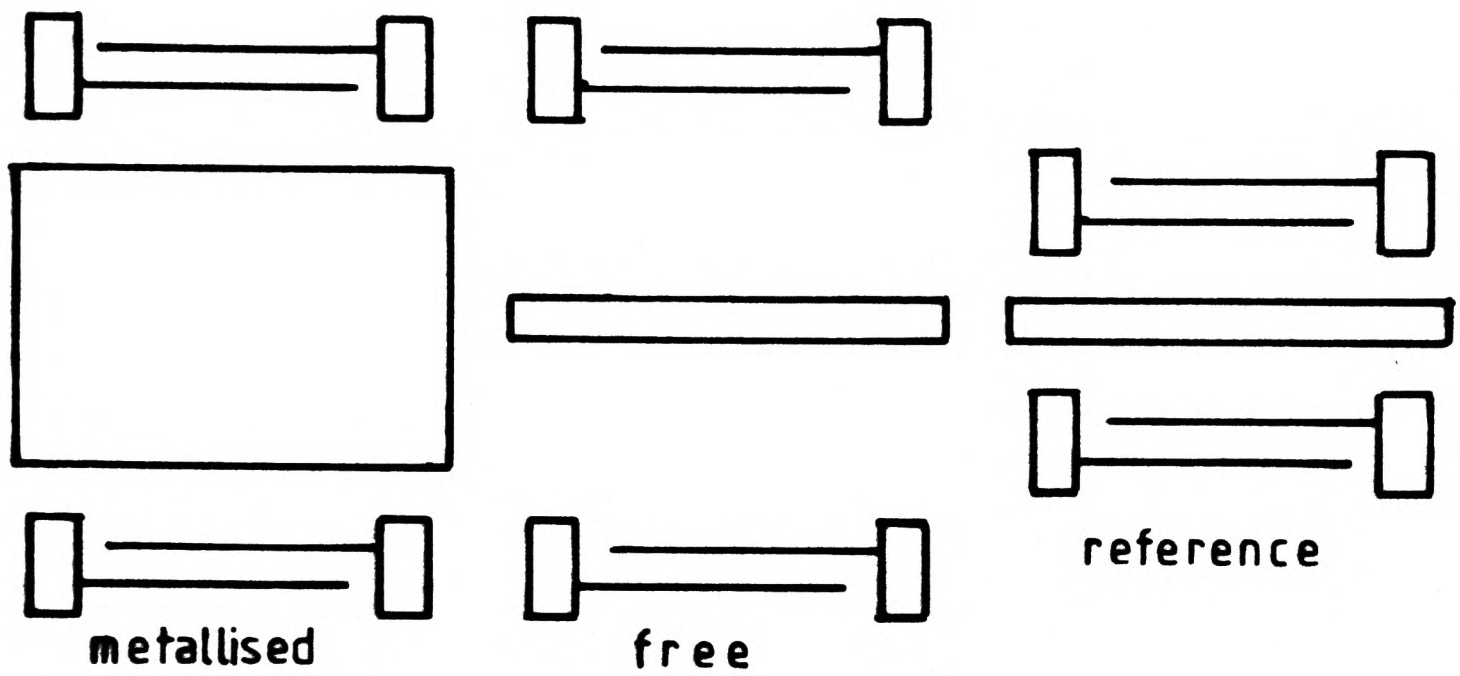


Fig. 2.14: Device to measure LSW velocity.

2.9 Radiation Resistance of Transducers in Various Directions

It is useful to know about LSW and bulk generation by arrays in various directions. A qualitative idea of this can be found by measuring the radiation resistance of transducers orientated in various directions. This was performed for 0°(Z), 16°, 45°, 74°, and 90°(X) propagating transducers using devices fabricated from existing masks. The transducer periods ranged from 50 to 70 microns, and the apertures from 4.2 to 12 mm.

The measurements were made with a calculator-controlled network analyser (Hewlett Packard 9825 and 8754A, respectively). The electrical signal was applied to the transducer with an HP 8660C signal generator, also calculator controlled, and the reflected electrical signal was measured. The radiation resistance was calculated using transmission line theory.* Software was developed by R. C. Woods and N. G. Doe.

* Transmission line theory gives the value of a load impedance (in this case, the transducer) at the end of a transmission line (the feed cable) in terms of the reflection coefficient.

CHAPTER 3

THEORY

The generation of SAWs by transducers has been discussed by several authors including Milsom [27,28] and Morgan [30]. In one approach, the electric charge on the transducer is taken to be the source of the SAW. In this chapter, this concept is carried over to calculate the reflectivity of dots. The electric potential associated with the SAW acts on the charge on the dot. The charge on a dot illuminated by a SAW is first calculated assuming elementary electrostatics, that is, neglecting piezo-electricity when considering the interactions between charges. The SAW "generated" by the illuminated dot is the reflected or scattered wave. A modification in the equations is required, because the charge on a dot varies in two dimensions, whereas the variation on a transducer is usually one-dimemsional, neglecting small end-effects. The theory is further modified to calculate velocity change in the dot array, and to find the reflectivity of holes in a thin metal film.

3.1 The Generation of SAWs by Electric Charge Varying in One Dimension

An accurate analysis for the generation of SAWs by a charge distribution varying in one dimension is given in ref. [27]. It is complicated by the facts that both mechanical and electrical relations are involved in a piezo-electric material, and that all piezo-electric materials are anisotropic. However, the mechanical and electrical properties of a piezoelectric solid can be reduced to a purely electrical problem by defining an effective permittivity ϵ_s such that

$$\bar{V}(k) = \frac{1}{|k|\ \epsilon_s(\frac{k}{\omega})}\ \bar{\sigma}(k) \qquad (3.1)$$

where $\bar{V}(k)$, $\bar{\sigma}(k)$ are respectively the spatial Fourier transforms of the charge at the surface of the substrate, and the resulting potential at the surface. Time variation of $\exp(j\omega t)$ will be supressed throughout. ϵ_s is a function of slowness s $(=k/\omega)$, and is computed numerically from the mechanical, electrical and piezo-electric properties of the particular material, thus embodying all the relevant information. ϵ_s for YZ $LiNbO_3$ is illustrated is fig. 3.1 (from ref. [27]) for positive values of s, being symmetrical about $s=0$. For a non-piezoelectric, isotropic material, ϵ_s is constant and equal to $(\epsilon_o+\epsilon)$. (This is shown in section 7 of this chapter. The additional term ϵ_o is due to the fact that the charge resides on the interface between solid and vacuum; otherwise 2ϵ would have been appropriate.)

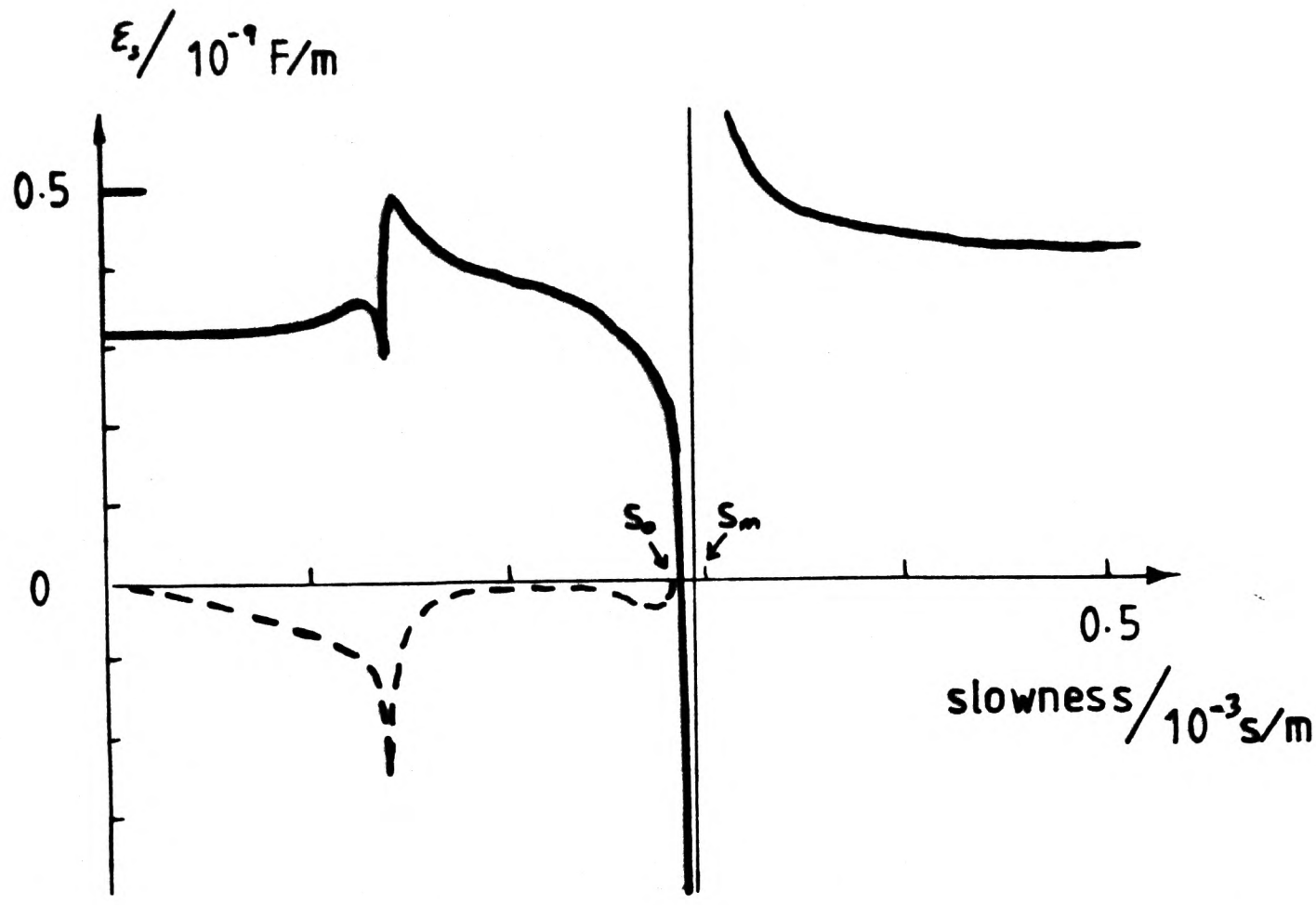


Fig. 3.1: Effective permitivity for Z-propagating wave on Y-cut LiNbO3. ___Real part, ---Imag. part.

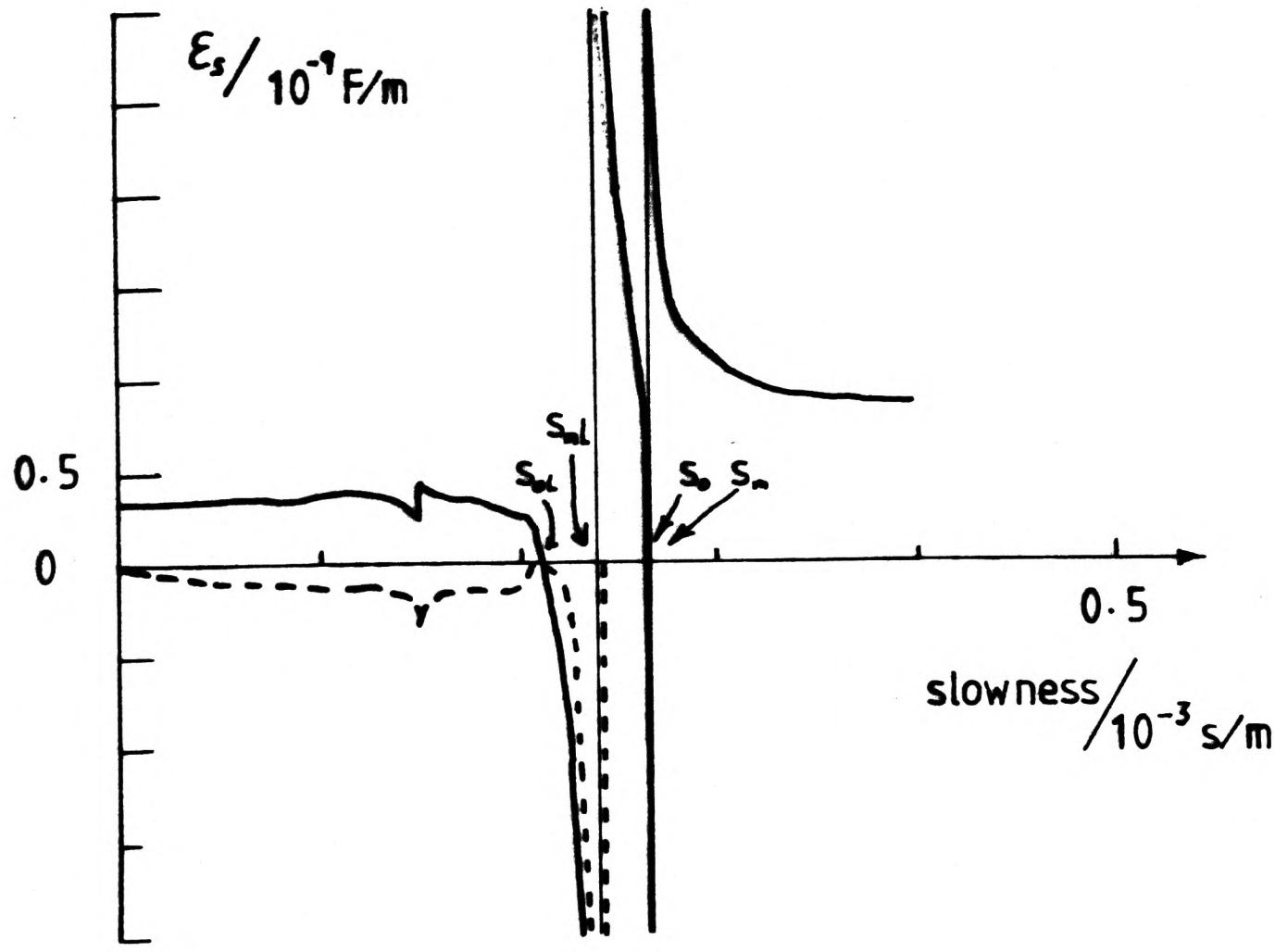


Fig. 3.2: Effective permitivity for X-propagating wave on Y-cut LiNbO3. ___Real part, ---Imag. part.

Applying the convolution theorem,

$$V(x) = \int_{-\infty}^{\infty} G(x-x') \, \sigma(x') \, dx' \qquad (3.2)$$

where the Green's function G(x) is given by

$$G(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\exp(-jkx)}{|k| \; \varepsilon_s(\frac{k}{\omega})} \, dk$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{|s| \; \varepsilon_s(s)} \exp(-j\omega s x) \, ds \qquad (3.2a)$$

So once the function ϵ_s has been computed, the potential at any point on the surface of the material can be found given the charge density. This potential is the sum of the electrostatic field due to the charge and the potential associated with the SAW and other waves generated by the charge.

The factor $\Gamma(s) = 1/|s|\epsilon_s(s)$ in the Green's function is more convenient than $1/|k|\epsilon_s(k/\omega)$ used earlier because it is independent of frequency. A method of calculating the Green's function accurately from $1/|s|\epsilon_s(s)$ is given in the appendix of reference [27]. The outline of a simplified version very similar to Morgan's [30] is used here. Inspection of ϵ_s in fig. 3.1 shows that $1/|s|\epsilon_s(s)$ has poles at $s=0$ and $s=\pm s_o$, s_o being the Rayleigh-wave slowness for a free surface. A function which has this form is

$$\frac{1}{|s|\ \varepsilon_s(s)} = \frac{1}{|s|\ \varepsilon_\infty} + \frac{2\, s_o\, G}{s^2 - s_o^2}$$

or

$$\Gamma(s) = \Gamma_e(s) + \Gamma_R(s) \qquad (3.3)$$

where $\Gamma_e(s)$ and $\Gamma_R(s)$ are respectively the electrostatic and Rayleigh-wave (SAW) contributions to $\Gamma(s)$, and G and $\epsilon_s(\infty)$ (abbreviated to ϵ_∞) are constants to be determined. Further justification for this function is given in reference [30]; it is accurate if bulk and leaky surface waves can be neglected in the particular material.

When $s = s_m$, that is, $k = k_m$, the wave number for a metallised surface, $\bar{\sigma}(k)$ is non-zero while $\bar{\Phi}(k)$ is zero, so, from (3.1),

$$\frac{1}{s_m\ \varepsilon_s(s_m)} = 0 \qquad (3.4)$$

or, from (3.3),

$$\frac{1}{s_m\ \varepsilon_\infty} + \frac{2\, s_o\, G}{s_m^2 - s_o^2} = 0 \qquad (3.5)$$

From (3.5), it can be shown that $G\epsilon_\infty \simeq \frac{v_m - v_o}{v_o} = \frac{\Delta v}{v}$, where $v = 1/s$. (This is done in Appendix 1 of Reference [30].) $G\epsilon_\infty$ is thus calculated from accurate empirical velocity data. (G and ϵ_∞ will not be required separately in what follows.)

Using methods similar to the appendix of ref. [27], the Green's function (3.2a) can now be evaluated in terms of ϵ_∞ and $\frac{\Delta v}{v}$.

$$G(x) = \frac{1}{-\pi \epsilon_\infty} \ln|x| - j G \exp(-jk_o|x|) \qquad (3.6)$$

where $G = (\frac{\Delta v}{v})\frac{1}{\epsilon_\infty}$. So (3.2) becomes

$$V(x) = \int_{-\infty}^{\infty} \frac{1}{-\pi\epsilon_\infty} \ln|x-x'| \sigma(x') dx' + \int_{-\infty}^{\infty} -j G \sigma(x') \exp(-jk_o|x-x'|) dx' \qquad (3.7)$$

The first term is the electrostatic contribution to potential. Provided that overall charge is zero, it can be shown that it decays with distance, so that far away from the transducer, it can be neglected. The second term relates the potential associated with the SAW to the generating electric charge. Its magnitude is constant with distance. At an observation point **x** to the right of all points **x'** with non-zero charge (**x>x'** for all **x'** such that $\sigma(x') \neq 0$), the modulus sign in the second term is unnecessary, so from (3.7),

$$V(x) = -j G \int_{-\infty}^{\infty} \sigma(x') \exp(-jk_o[x-x']) dx' \qquad (3.8)$$

$$V(x) = -j G \exp(-jk_o x) \bar{\sigma}(-k_o) \qquad (3.9)$$

where $\bar{\sigma}(k)$ is the spatial Fourier transform of $\sigma(x)$. It is apparent from (3.8) that **V(x)** is obtained by multiplying each element of charge by an appropriate phase delay, summing, and inserting a constant of proportionality.

3.2 The Effect of the LSW

The foregoing analysis has been found adequate for transducers on **Z** propagating Y-cut $LiNbO_3$, and on some other materials. However, when is was applied to dot arrays, in

which propagation in the **X** direction was involved, reflectivities predicted were larger than those measured, by a factor of approximately 1.4 . The above analysis was modified as a part of this research project.

The discrepancy is due to the Leaky Surface Wave (LSW), which has very strong piezoelectric coupling in the **X** direction, but zero coupling in **Z** [21]. The leaky wave, because of its very high coupling, has a significant effect on the relationship between $G\epsilon_\infty$ and $(v_m - v_o)/v_o$. The value of $G\epsilon_\infty$ has been calculated numerically by Milsom [29]. Morgan's simplified method [30] will now be extended to provide an alternative calculation for $G\epsilon_\infty$. Compared with $\epsilon_s(s)$ for the **Z** direction, the effective permittivity $\epsilon_s(s)$ for **X** (fig. 3.2, [29]) has an additional pole and zero, corresponding to the LSW. There thus is an additional term in the Green's function, which is assumed to take the same from as the surface wave term. (3.3) is modified to:

$$\frac{1}{|s|\,\varepsilon_s(s)} = \frac{1}{|s|\,\varepsilon_\infty} + \frac{2\,s_o\,G}{s^2 - s_o^2} + \frac{2\,s_{oL}\,G_L}{s^2 - s_{oL}^2} \qquad (3.10)$$

$1/|s|\epsilon_s(s)$ is zero at two values of s, s_m , the metallised-surface SAW slowness, and s_{mL}, the metallised-surface leaky-wave slowness. It is thus possible to form two simultaneous equations to find $G\epsilon_\infty$ (and $G_L\epsilon_\infty$, if required). The effect of the new term in (3.10) is significant in the **X** crystal direction, but not in **Z** (fig. 3.10).

3.3 Reflections and Scattering by Widely-Separated Dots

The main part of the theoretical work on dots in this research project will now be described. The analysis for transducers is extended for use with dots. The dots studied are rectangular, with sides inclined at about 45° to the Z axis.

3.3.1 Charge on a dot

The first stage in the analysis to calculate reflectivity is to compute the charge distribution inside a dot illuminated by a SAW. For a one-dimensional array of strips, this can be done analytically using a Fourier series whose coefficients are Legendre polynomials [9]. For the dots, in which charge varies in two dimensions, a numerical technique was chosen. Piezo-electric coupling is assumed to be relatively weak, so that, at close range (distances of the order of the dot length, usually less than a couple of wavelengths), the electrostatic term in the Green's function dominates. This condition is satisfied by most piezoelectric materials, including $LiNbO_3$. The charge distribution can then be found using simple electrostatics. The permittivity ϵ of the material is assumed to be isotropic, although the results are not changed by this assumption, as shown in section 3.4 . Since elementary electrostatics is assumed, elastic properties, isotropic or otherwise, are not relevant. The dots are assumed to be sufficiently widely-separated so as not to interact.

In the absence of dots, a piezoelectric wave travelling in the positive-z direction has a potential at the surface of $Re[V_o \exp j(-k_o z+\omega t)]$ associated with it. The time variation, and the symbol Re[], will be suppressed throughout. Redistribution of charge on the dot causes the potential there to be at a uniform, floating value V_d (fig.3.3) , which is computed together with the charge density $\sigma(z,x)$. Inside the dot, the change ΔV in potential due to the charge alone is $V_d - V_o \exp(-jk_o z)$. Just outside, it has some value between this and zero; this value is not required in the computation.

Assuming elementary electrostatics,the potential ΔV at a distance l from a point charge q is

$$\Delta V = \frac{1}{2\pi(\varepsilon_o+\varepsilon)} \frac{q}{l} \qquad (3.11)$$

where $2\pi(\epsilon_o+\epsilon)$ replaces $4\pi\epsilon_o$ because the charge is on the boundary between vacuum and the piezelectric material. The dot is divided into n unequal-sized "cells" (fig. 3.4), each of which contains a charge q_j. The cells are smaller near the edges of the dot because charge density there is both large and more rapidly-varying. Assuming for the moment that q_j is lumped at the centre of each cell, potential ΔV_i at the centre of any cell j is given by

$$\Delta V_i = V_d - V_o \exp(-jk_o z_i) = \frac{1}{2\pi(\varepsilon_o+\varepsilon)} \sum_{j=1}^{n} \frac{1}{l_{ij}} q_j \qquad (3.12)$$

l_{ij} is the distance between the centres of the ith and jth cells. The terms $(1/l_{ij})$ when $l_{ij}=0$ have to be modified as described later. Since the dot is isolated,

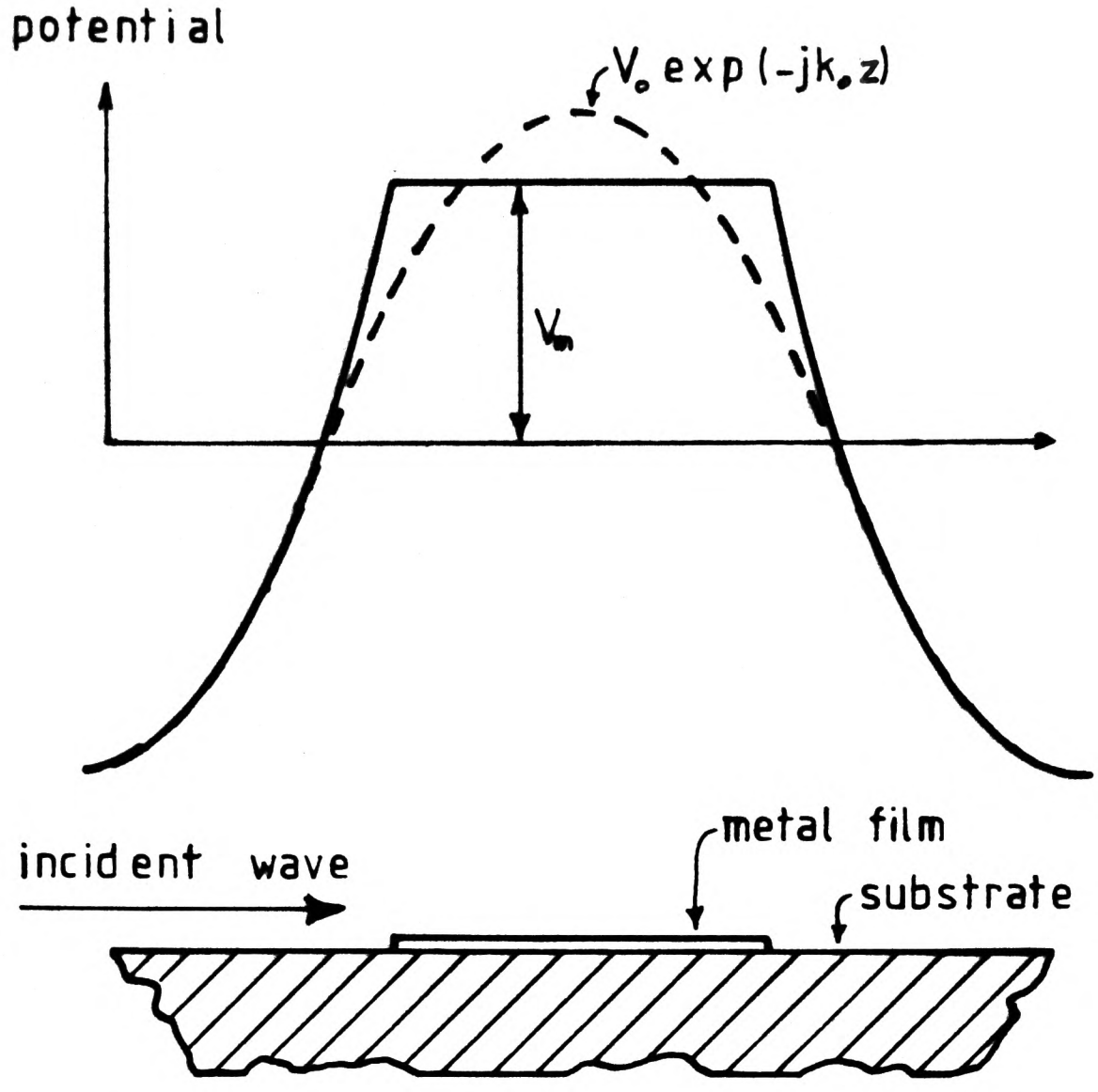


Fig. 3.3: Electric potential at the surface on and near a dot.

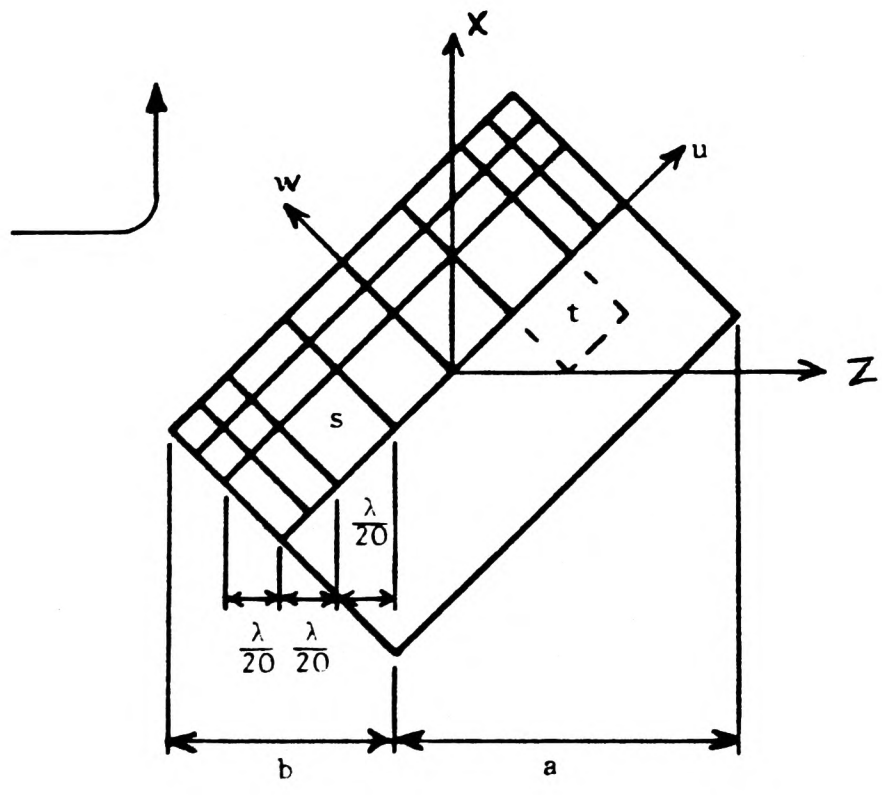


Fig. 3.4: Unequal-sized cells into which a dot is divided for numerical computation.

$$\sum_{j=1}^{n} q_j = 0 \qquad (3.13)$$

(3.12) and (3.13) form (n+1) simultaneous equations, from which V_d and n values of q_j can be found. (3.12) can be written

$$\left.\begin{aligned} 1 &= \frac{1}{2\pi(\varepsilon_o+\varepsilon)} M_{ij}\, h_j \\ -V_o \exp(-jk_o z_i) &= \frac{1}{2\pi(\varepsilon_o+\varepsilon)} M_{ij}\, q_j' \end{aligned}\right\} \qquad (3.14)$$

where

$$q_j = V_d h_j + q_j' \qquad (3.15)$$

M_{ij} is a matrix with terms of the form $(1/l_{ij})$. Solving either set of simultaneous equations for h_j and q_j' is equivalent to inverting (nxn) matrix $\underline{M}$, which is common to both sets. Using (3.13) and (3.15) gives

$$V_d = -\sum_j q_j' \Big/ \sum_j h_j \qquad (3.16)$$

after which n values of q_j are found from (3.15).

Equation (3.12) has to be modified before it is used. When i=j, $(1/l_{ij})$ is infinite. This is a consequence of assuming that charge is lumped at the centre of each cell. A more accurate approximation is to assume that charge is uniformly distributed over the cell. For this purpose, it was convenient to divide each of the largest cells (near the centre of the dot, fig. 3.4) into 4 sub-cells, each with equal charge $q_j/4$, and the narrow cells into two. All

sub-cells so defined have the same size. The number of unknowns q_j is not increased. It can be shown that in fig. 3.5,

$$\Delta V_1 \simeq \frac{1}{2\pi(\epsilon_o+\epsilon)}\left\{ \frac{1}{l_o}\left(\frac{1}{0.2836}\right)q_1 + \frac{1}{l_2}\left(1+0.04\left(\frac{l_o}{l_2}\right)^2\right)\frac{q_2}{2} + \frac{1}{l_2'}\left(1+0.04\left(\frac{l_o}{l_2'}\right)^2\right)\frac{q_2}{2} + \text{-- --}\right\} \quad (3.17)$$

The first coefficient $1/0.2836l_o$ is found by direct integration, while the other coefficients of the form

$$\frac{1}{l}\left(1+0.04\left(\frac{l_o}{l}\right)^2\right)$$

are found to be a suitable approximation by comparison with the exact values, also found by integration.

A second modification was also made. It is clear from symmetry conditions that $\Delta V_s, \Delta V_t$, and q_s, q_t (fig. 3.4) are complex conjugates. Accordingly, real and imaginary parts of q_j were found separately, using $Re(\Delta V_i)$ and $Im(\Delta V_i)$ respectively. Since $Re(q_s)=Re(q_t)$ and $Im(q_s)=-Im(q_t)$, the number of unknowns was halved and the computation time reduced by a factor of eight.

3.3.2 Reflections through 90°

Unlike most transducers, a dot has a charge distribution which varies in two dimensions. The second stage of the model requires a two-dimensional version of equation (3.9).

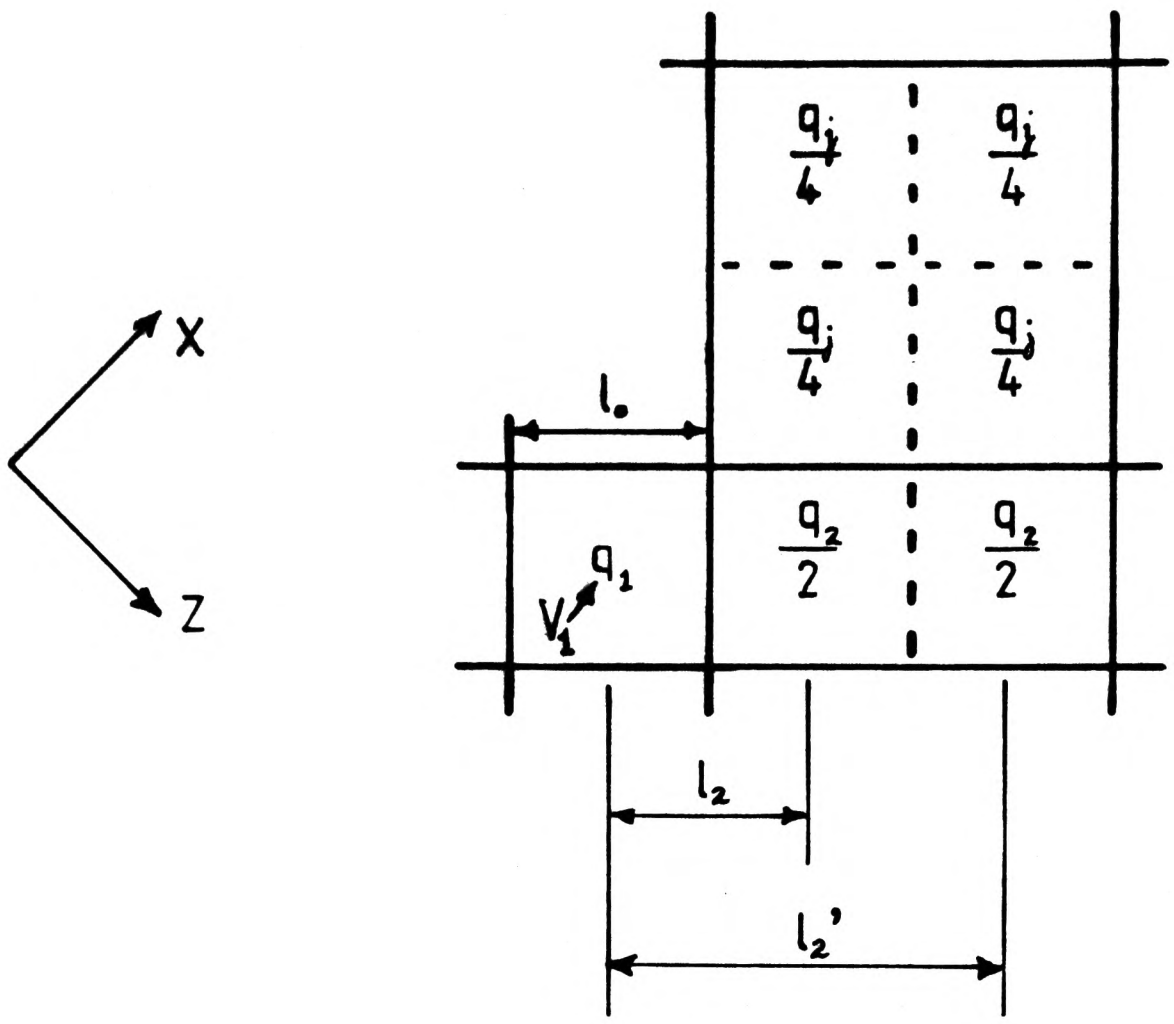


Fig. 3.5: Charge distribution assumed in computations.

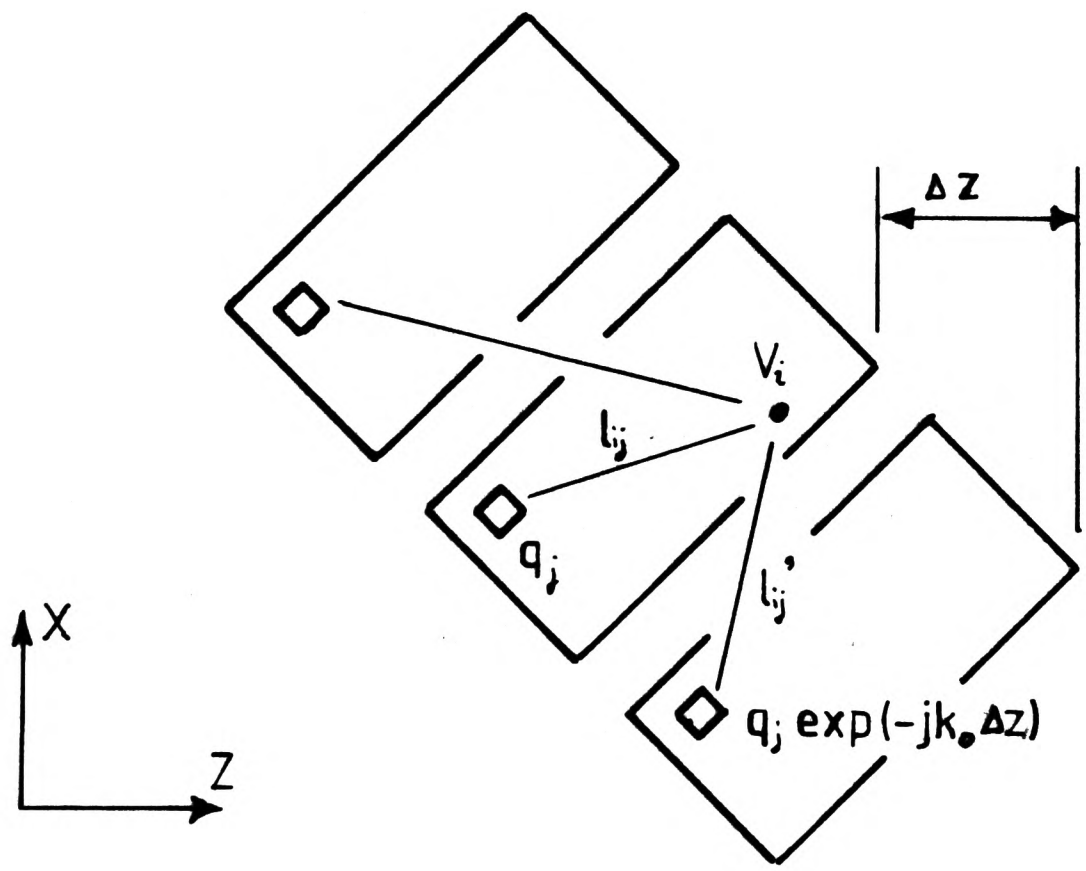


Fig. 3.6: Accounting for charge on neighbouring dots.

The amplitude of the **X** propagating SAW is required at a distance **D** from an infinitely long line of dots. **D** is taken to be much greater than the separation between dots L_1 ($D >> L_1$). The configuration is illustrated in fig. 3.7 (in which the non-standard **Z-X** axes should be noted). It is assumed that the dots are in phase with each other. $\sigma(x)$, constant with **z** in the one-dimensional case, is replaced by its average value

$$\frac{1}{L_1}\int_{-L_1/2}^{L_1/2} \sigma(z,x)\, dz$$

in the two dimensional dot array. At distance **D**, variations of $\sigma(z,x)$ along **z** have no effect; it is this average value which is relevant. Substituting this into (3.9) gives

$$V_r(x) = -j\, G_x \exp(-jk_o x)\, \frac{\bar{\sigma}(0,-k_o)}{L_1} \qquad (3.18)$$

where $\bar{\sigma}(k_z,k_x)$ is the two-dimensional spatial Fourier transform of the charge density on one dot. The amplitude of the reflected wave from a row of dots in fig. 3.8 is also given by (3.18), except that the dots do not have to be in phase, since the path length between an incident and a reflected wavefront is always constant for all dots in a suitably orientated row.

$V_r(0)/V_o$ gives the reflectivity of a row of dots in terms of a ratio of potentials. On an anisotropic material, this is not equal to the square root of the ratio of power densities, which is a more useful quantity. Accordingly, reflectivity jr is defined such that its magnitude is $\sqrt{(P_r/P_o)}$, and its phase is equal to the phase of V_r/V_o. To find jr, two reflections

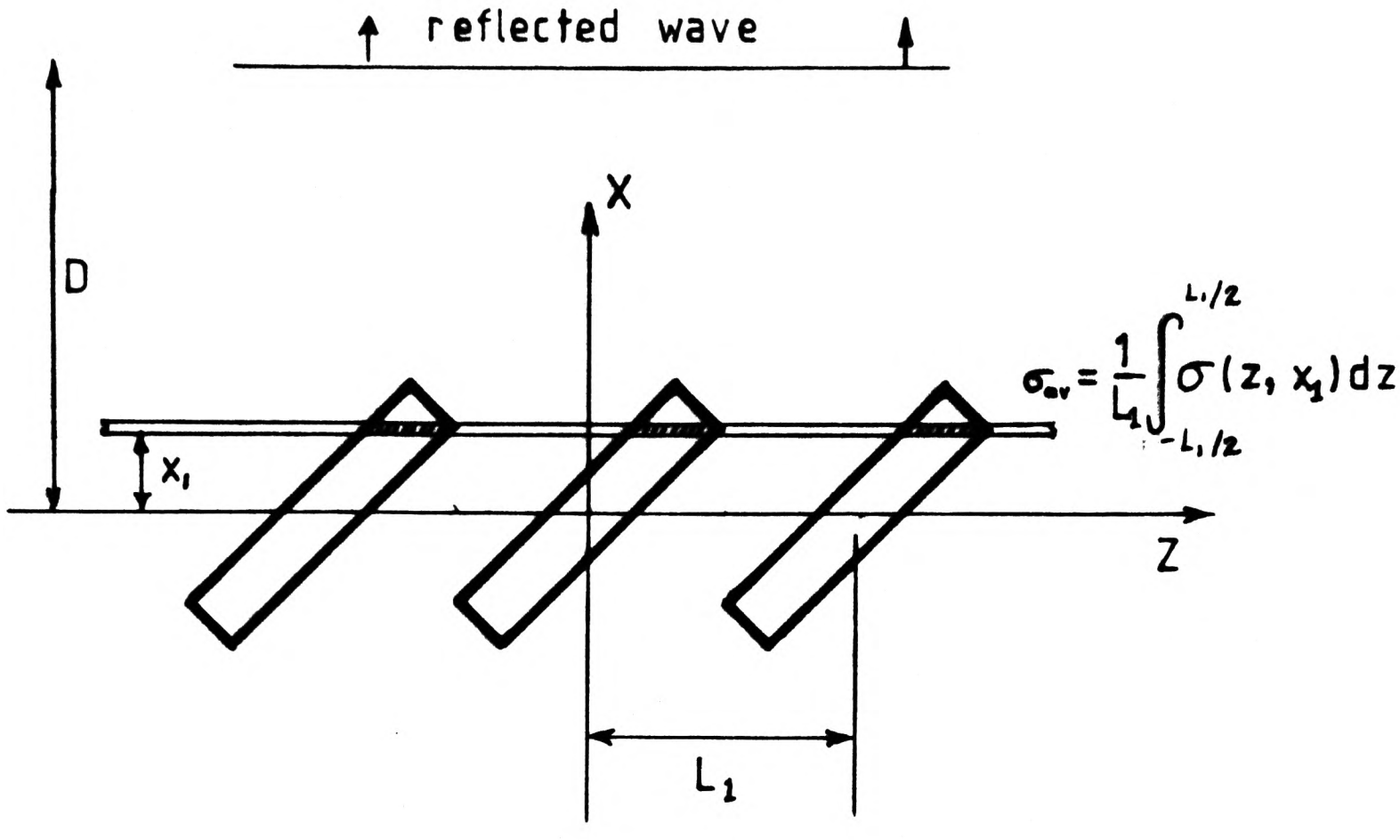


Fig. 3.7: Computation of reflected wave from charge density.

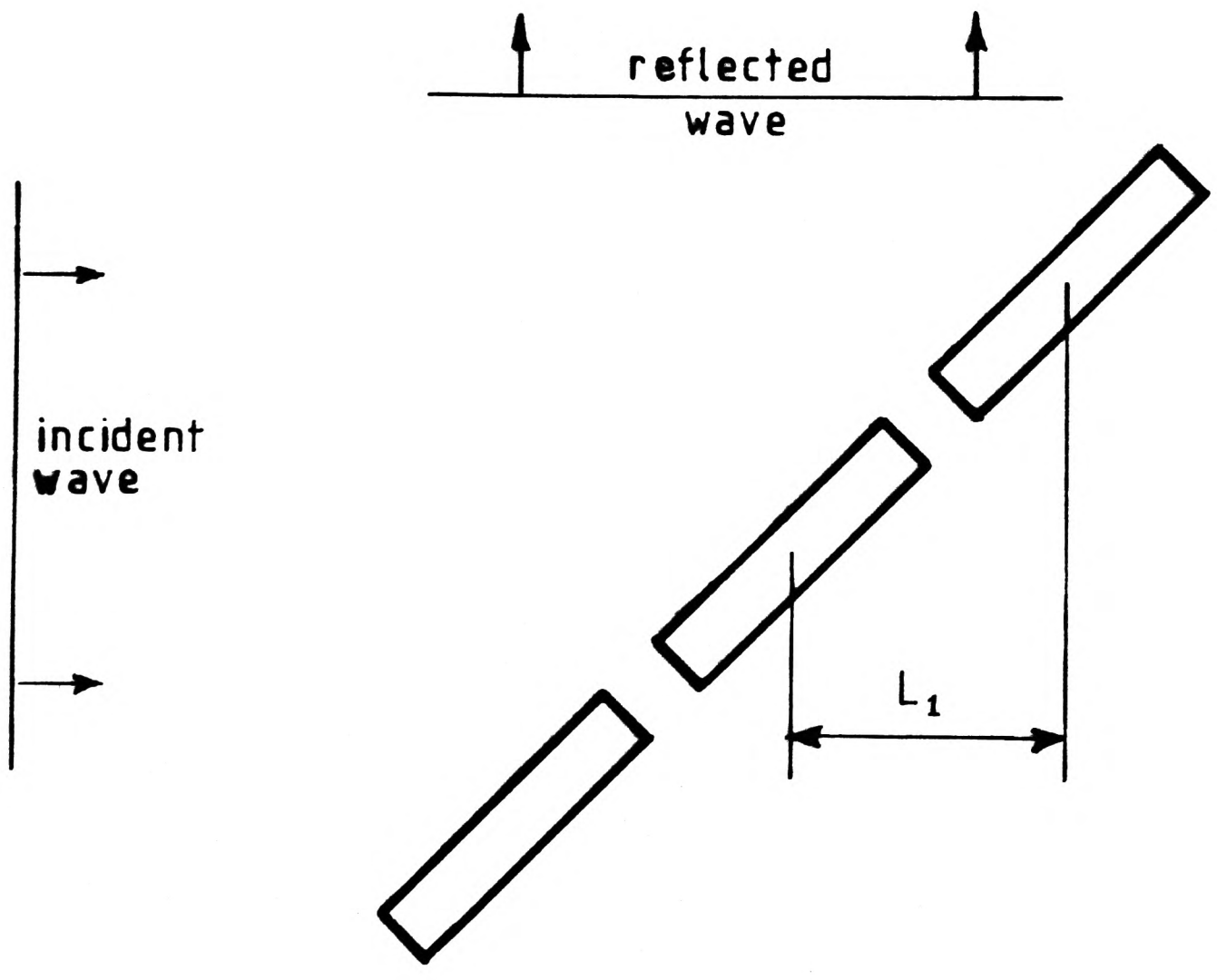


Fig. 3.8: Reflection from a "row" of dots.

by two rows of identical dots are considered, so that the propagation direction changes from **Z** to **X** to **Z**. The final output is

$$j^2 G_z G_x \left(\frac{\bar{\sigma}(0,-k_o)}{L_1}\right)^2$$

so that in a single reflection,

$$jr = -j\sqrt{G_z \epsilon_\infty G_x \epsilon_\infty}\ \frac{\bar{\sigma}(0,-k_o)}{L_1 \epsilon_\infty V_o} \qquad (3.19)$$

where $(G_z \epsilon_\infty)$ and $(G_x \epsilon_\infty)$ are constants which are determined as described in sections 1 and 2 of this chapter. There is a factor $(\epsilon_o+\epsilon)$ implicit in $\bar{\sigma}(0,-k_o)$, which, for a weakly piezo-elctric material, cancels with ϵ_∞ in the denominator [30]. $\bar{\sigma}(0,-k_o)$ also contains a factor V_o, which also cancels with a term in the denominator.

3.3.3 Reflections by Dots in Other Directions

Constructive interference between dots in a row may occur for additional propagation directions, if the path difference is an integer number of wavelengths. Most of these are undesirable. In this thesis, this has been termed "scattering". Using arguements similar to the previous section, the reflectivity of the row in the direction (k_z, k_x), where $\sqrt{(k_z^2+k_x^2)} = k_o$, is

$$jr = -j\sqrt{G_z G}\ \frac{\bar{\sigma}(-k_z,-k_x)}{L V_o} \qquad (3.20)$$

where **G** is a function of direction. **L** is the separation between dots measured as a projection on the reflected wavefront, and therefore also varies with direction.

The computer program calculates **r** for 0°, 90°, -90° and 180° reflections for a range of dot "lengths" **a** (fig. 3.4), 0.1 to 1.0 **λ**, and "widths" **b**, 0.1, 0.2, 0.3 and 0.4 **λ**. The time-consuming part of the computation, finding charge density, only has to be done once for each size of dot. Finding the Fourier transform $\bar{\sigma}(k_z,k_x)$ takes an insignificant amount of time as it has to be computed only for a few values of (k_z,k_x) . The program takes about 25 minutes of CPU time.

3.3.4 Anisotropy

The anisotropy of the mechanical properties of the substrate are irrelevant to the calculation of charge density, since electrostatics is assumed. In the second stage of the analysis, in which the reflected wave is calculated, the mechanical properties determine **G** and k_o as already described; they are therefore functions of direction.

Both stages of the calculation assume that permittivity ϵ is isotropic. Further computations were made for anisotropic media using the following arguements. On an anisotropic material, Poisson's equation is modified to

$$\sum \varepsilon_{ij} \frac{\partial^2 V}{\partial x_i \partial x_j} = -\rho \qquad (3.21)$$

for $LiNbO_3$,

$$\varepsilon_{xx} = 44\varepsilon_o$$

$$\varepsilon_{yy} = 44\varepsilon_o$$

$$\varepsilon_{zz} = 29\varepsilon_o$$

$$\varepsilon_{ij\,(i\neq j)} = 0 \qquad (3.22)$$

Substituting $\epsilon=\epsilon_{xx}$, $\eta=\sqrt{(44/29)}$, $w=\eta z$,

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial w^2} \right) V = - \frac{\rho'}{\varepsilon} \eta \qquad (3.23)$$

where ρ' is the limit of the value

$$\frac{\Delta q}{\Delta x \, \Delta y \, \Delta w}$$

as the quantities in the denominator tend to zero. (3.23) shows that Poisson's equation can be applied to $LiNbO_3$ by suitable scaling of **z** and **q**. Gauss's theorem, used to derive (3.11), can be obtained from Poisson's

A reflectivity program was written with these modifications, for reflections through 90°. There were no significant deviations from the previous results, so the question of anisotropy of ϵ has not been considered further.

3.3.5 Further Computations

A few further computations were made by modifying the original program. Experimental data for dots with width **b**=0.133, 0.267 and 0.4 wavelengths is available. Matching computed data is obtained by changing the ratio of wavelength to cell size so that the largest cell (fig. 3.4) measures $\lambda/15$ instead of $\lambda/20$.

A second modification was to increase the number of cells, halving their linear dimensions, to obtain more accurate values of **r**. The number of cells is increased by a factor of about 4, so time taken increases by $4^3=64$; nine dot sizes consumed more that 2 hours of CPU time. The maximum deviation between these 90° results and the previous ones is

about 0.3×10^{-3}, or only about 0.02 of the maximum reflectivity calculated.

Relfectivity of long dots ($1<a/\lambda<2$, $0.1<b/\lambda<0.4$) was also computed. The results indicate that for $L_1=2a$ and **b** constant, **r** varies by less than 0.5×10^{-3} over this range of **a**.

3.4 Reflections and Scattering by Dots with Near-Neighbours

The foregoing analysis is now extended to include capacitative interactions between dots. The potential at any point in a dot is mainly determined by the incident SAW and the charge on the dot, as previously discussed, but partly determined by charge on neighbouring dots. The electrostatic or capacitative influence is short-ranged, particularly since overall charge on a dot is zero (as in a dipole). Only the neighbours within about one or two SAW wavelengths need to be considered. Although the dots also interact via multiple reflections of SAWs, these can be accounted for where necessary using standard numerical [31] or analytic [49] techniques for multiple reflections, not dealt with here. (These two types of interaction correspond to the two terms in the Green's function in equation (3.6).) It is therefore only necessary to consider localised capacitative interactions. The electrodes of a transducer, in contrast, interact via their electrical connection, so it is necessary to find a solution for all the electrodes simultaneously.

If a dot array is approximately periodic, that is, if the dot sizes and spacings are constant or vary very slowly, then for each charge q_j on a dot, there is a charge $q_j \exp(-jk_o\Delta z)$ at the corresponding position in the neighbouring dot (fig. 3.6). Δz is the increase in z-coordinate of the neighbouring dot. Equation (3.12) is modified to:

$$\Delta V_i = \frac{1}{2\pi(\varepsilon_o+\varepsilon)} \sum_{j=1}^{n} \left(\frac{1}{l_{ij}} + \frac{\exp(-jk_o\Delta z)}{l'_{ij}} + \text{.....} \right) q_j \qquad (3.24)$$

l_{ij} being shown in fig. 3.6 . The number of unknowns q_j is not increased by increasing the number of neighbours taken into consideration. However, since the coefficients of q_j are now complex, it is no longer convenient to calculate real and imaginary parts separately, nor to use the symmetry condition to halve the number of unknowns.

A FORTRAN program was written to compute scattering coefficients at 5° intervals, including 90° reflections and 0° forward scatttering, at 6 frequencies from $0.6f_o$ to $1.6f_o$, for any specified dot size ($a/\lambda_o, b/\lambda_o$) and specified positions of nearest neighbours. Normalised values of $L=\lambda_o$ and $\sqrt{(G_z G)}\epsilon=1$ were used in equation (3.20). The results are presented as a polar diagram (figs. 3.15,4.9)

The basis for a more useful presentation of the above data will now be discussed. This presentation allows the scattering coefficient per dot (the "element factor") already computed to be combined with the effect of the array pattern (the "array factor"). It is valid for periodic or gradually-varying arrays. An isotropic material is assumed. This can be extended to anisotropic media by replacing the

circles described below with the slowness curves of the material. (In a three dimensional analysis, the circles would be "Ewald spheres").

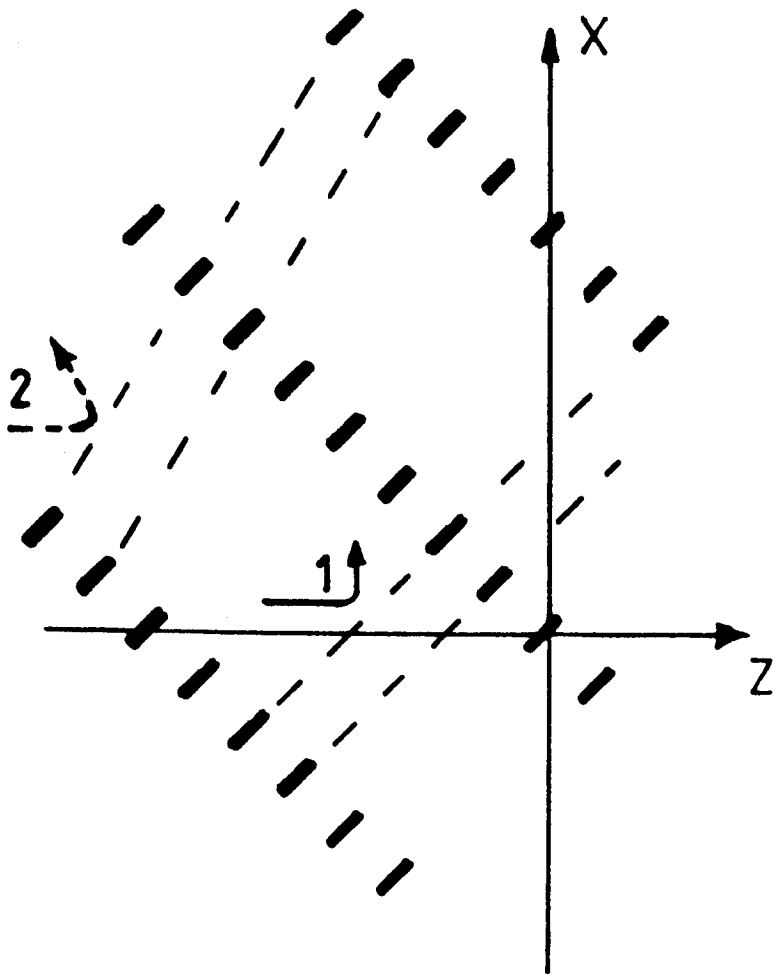
An infinite, periodic array of isotropic point scatterers (fig. 3.9(i)) has a spatial Fourier transform $H(k_z,k_x)$ which is zero everywhere except for a set of delta functions. $H(-k_z,-k_x)$ is shown in fig. 3.9(ii). In a finite array, each delta function is spread out over a small area. The vector joining the origin to one of the delta functions in $H(-k_z,-k_x)$ is described in reference [47] as a "scattering vector" $\underline{k}_g$. As explained in that reference, and in Appendix I, a reflected wave with wave vector $\underline{k}_{out}$ must satisfy the conditions

$$\underline{k}_{out} = \underline{k}_{in} + \underline{k}_g \qquad (3.25)$$

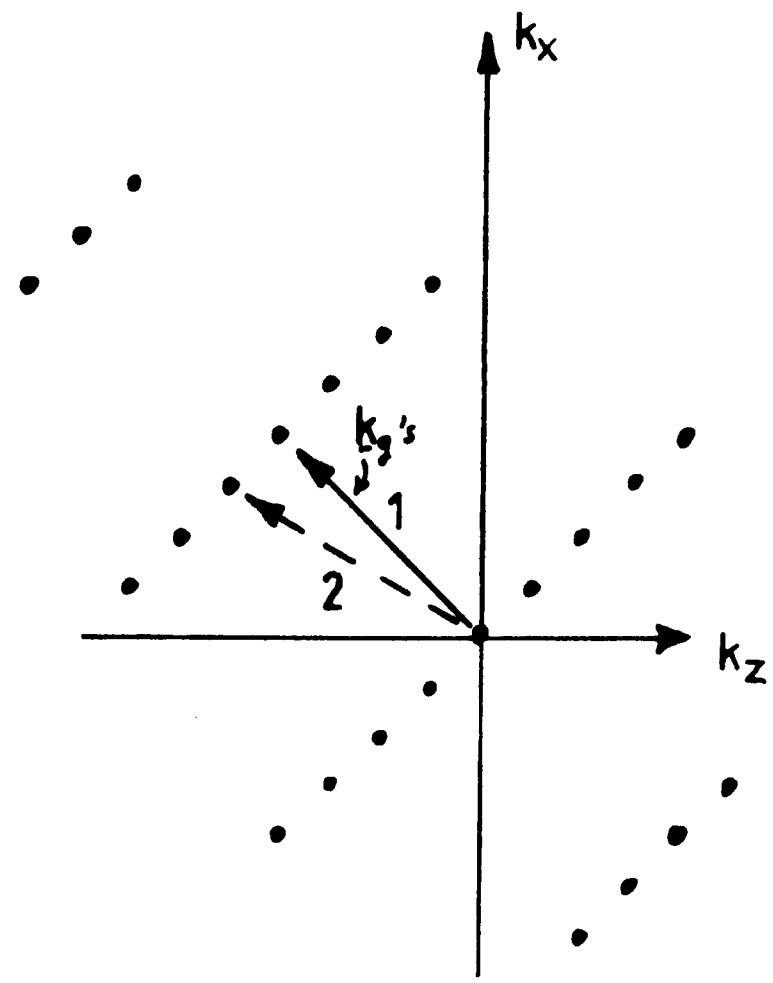
$$|\underline{k}_{out}| = |\underline{k}_{in}| \qquad (3.26)$$

They are satisfied for a particular frequency if one of the delta functions of the Fourier transform lies on the circle (fig. 3.9(iii)) whose radius is specified by $\underline{k}_{in}$. The amplitude of reflection is proportional to the delta function. In each array considered here, the delta functions are equal. The arrays intended for devices can be sparser than shown in fig. 3.9(i), so that the number of scattering vectors is large. Each scattering vector corresponds to a set of Bragg lines, that is, to a scattering mode (figs. 3.9(i) and (ii)).

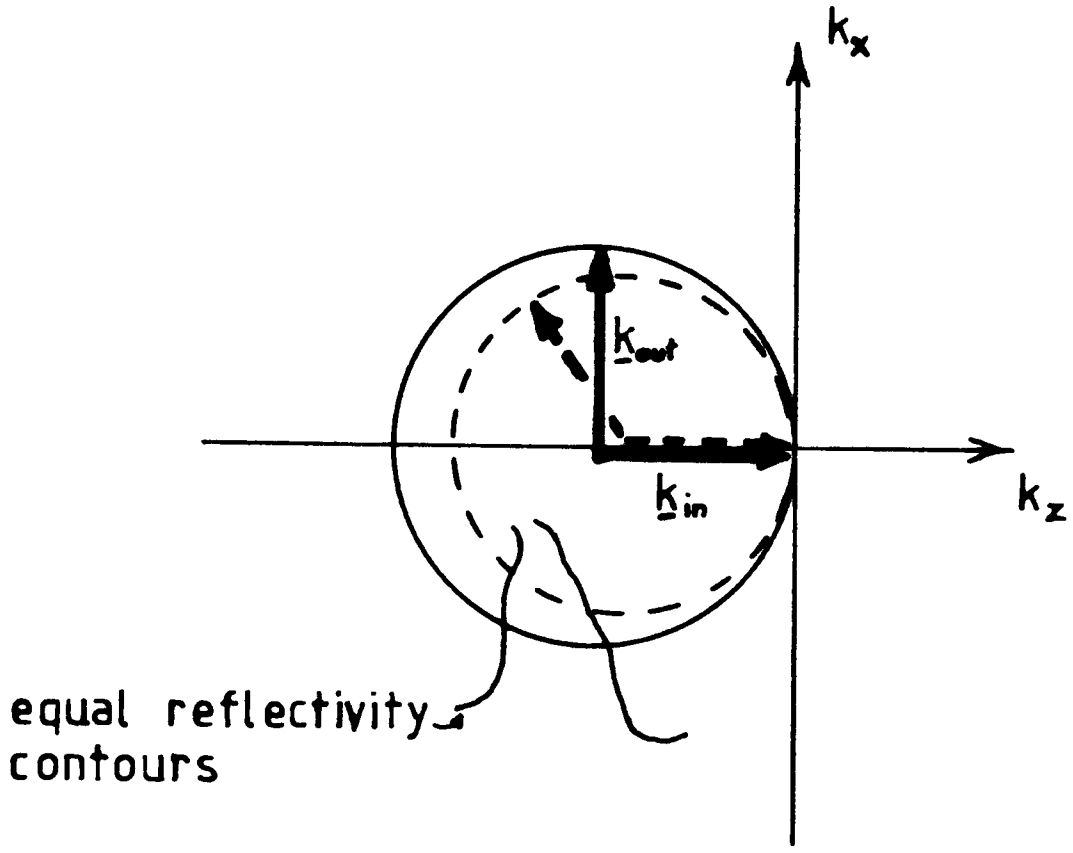
If the isotropic reflectors are replaced by dots, then the amplitude of reflection also depends on the scattering coefficient at that particular direction and frequency. It is



(i) Array pattern.



(ii) Spatial Fourier Transform of Array pattern.



(iii) Ewald circles, with Dot reflectivity contours superimposed.

Fig. 3.9:

convenient to map the scattering function $r(\theta, k_{in})$ onto the Ewald circles in fig. 3.9(iii), presenting **r** as a set of equal-amplitude contours. The amplitude associated with each scattering vector can then be read off the contour map.

The computer program uses the data already found. $G\epsilon_\infty$ is calculated as a function of direction as described in section 2 of this chapter using data from [1] and [21], and is shown in fig. (3.10). **L** remains normalised to $L=\lambda_o$. A standard contour plotting program, which interpolates between the given data, was modified to fit the non-rectangular grid (fig. 3.9(iii)) by N. G. Doe. The scattering coefficient associated with each scattering vector is read off from the contour diagram. If the array pattern (and hence the scattering vectors) is changed, the reflectivity contours are shifted only slightly because of the different positions of the nearest neighbours, so an approximate value can still be found from the same contour diagram.

3.5 Velocity Change

The transmitted wave after a row of dots can be considered to be made up of the incident wave plus the forward scattered wave, which is 90° out of phase. If the forward scattering coefficient is jr, where **r** is small, the dots change the phase of the transmitted wave by **r** radians per wavelength. If there is one row per wavelength along the propagation direction, the fractional change in phase velocity is $r/2\pi$. **r** as a function of frequency is already available from the first program described in section 4.

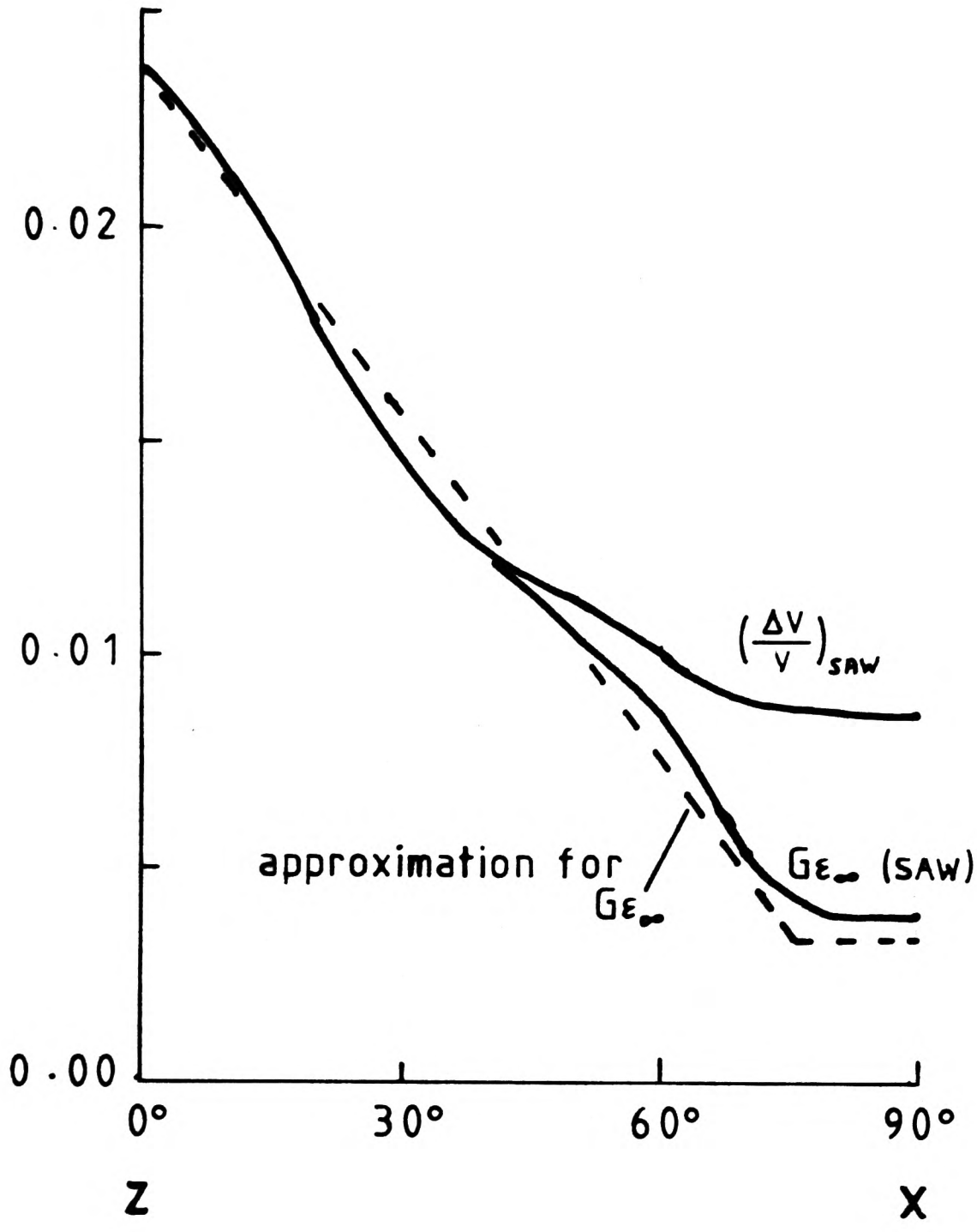


Fig. 3.10: $G\varepsilon_\infty$ as a function of direction on Y-cut LiNbO3.

3.6 90° Reflections by Holes

The numerical model is similar to the one used for dot reflections, but is slightly modified as follows in the next paragraph. The reason for making the modifications is that the metal film can cover a much larger area than the hole. If all the electric charge on the metal film were approximated by unknown charges q_j, solving for these charges would take a prohibitive amount of computer time.

Outside the metal film, the incident wave has a potential $V_o \exp(-jk_o z)$. Inside the film, the potential is exactly cancelled by the electrostatic field $-V_o \exp(-jk_o z)$ of the charge. From equation (3.39),

$$\sigma_o = -|k_o|(\varepsilon_o+\varepsilon)V_o \qquad (3.27)$$

q_o, the charge per "cell", is easily found from σ_o.

A completely metalised surface through which a SAW is travelling contains a sinusoidally-varying charge distribution $\sigma_o \exp(-jk_o z)$, as in fig. 3.11(i), but the overall reflection is zero. By the principle of superposition, it is only necessary to find the change in charge distribution (fig. 3.11(ii)) due to the hole. This consists of $-\sigma_o \exp(-jk_o z)$ inside the hole, together with the redistribution of charge around it. This redistribution is such that it restores every point on the metal film to ground potential. Thus although charge is distibuted all over the metal film, only a small region near and in the hole needs to be considered.

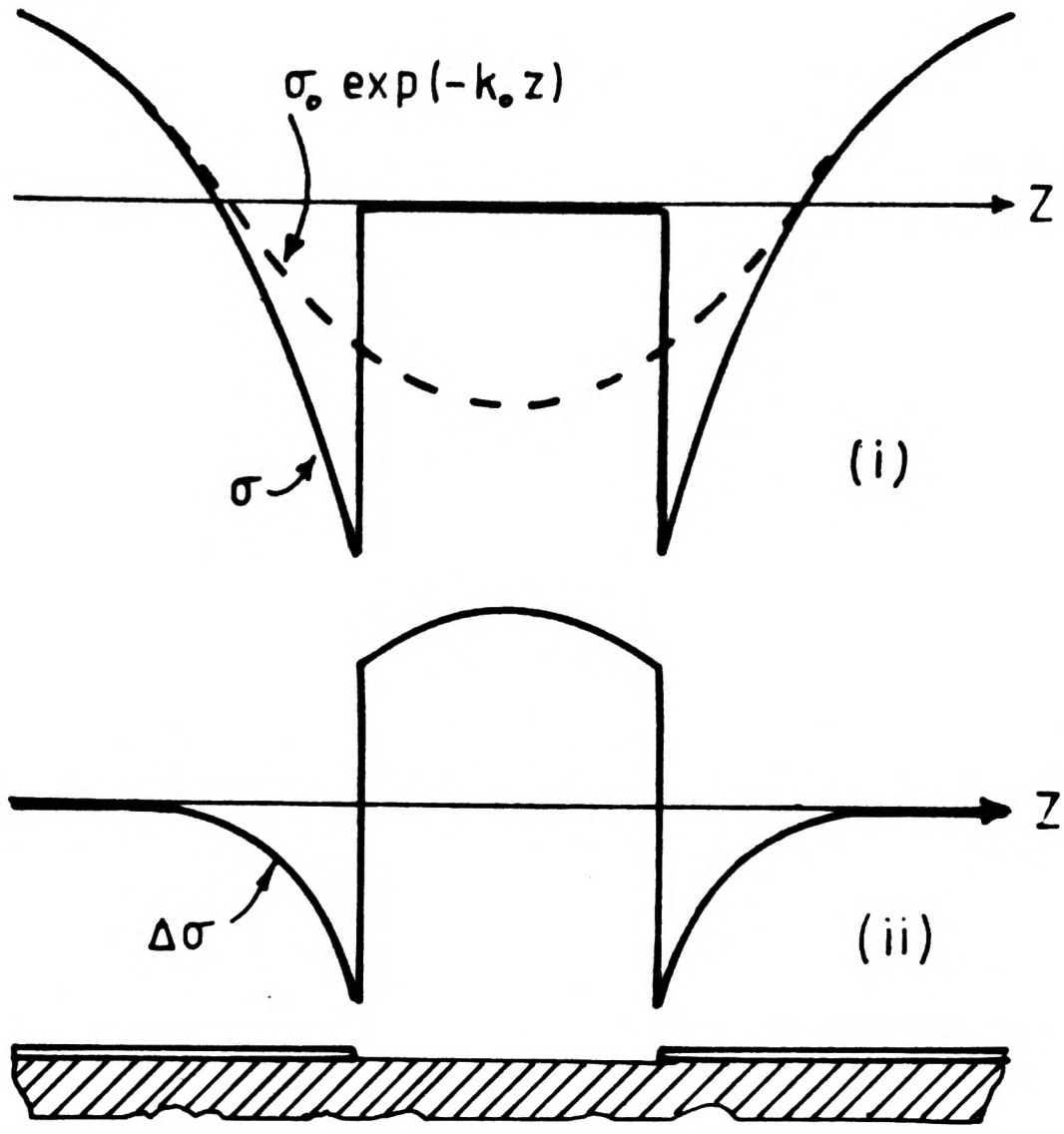


Fig. 3.11: (i) Charge density on metallised surface (dashed) and metallised surface with a hole.
(ii) Change in charge density due to the hole.

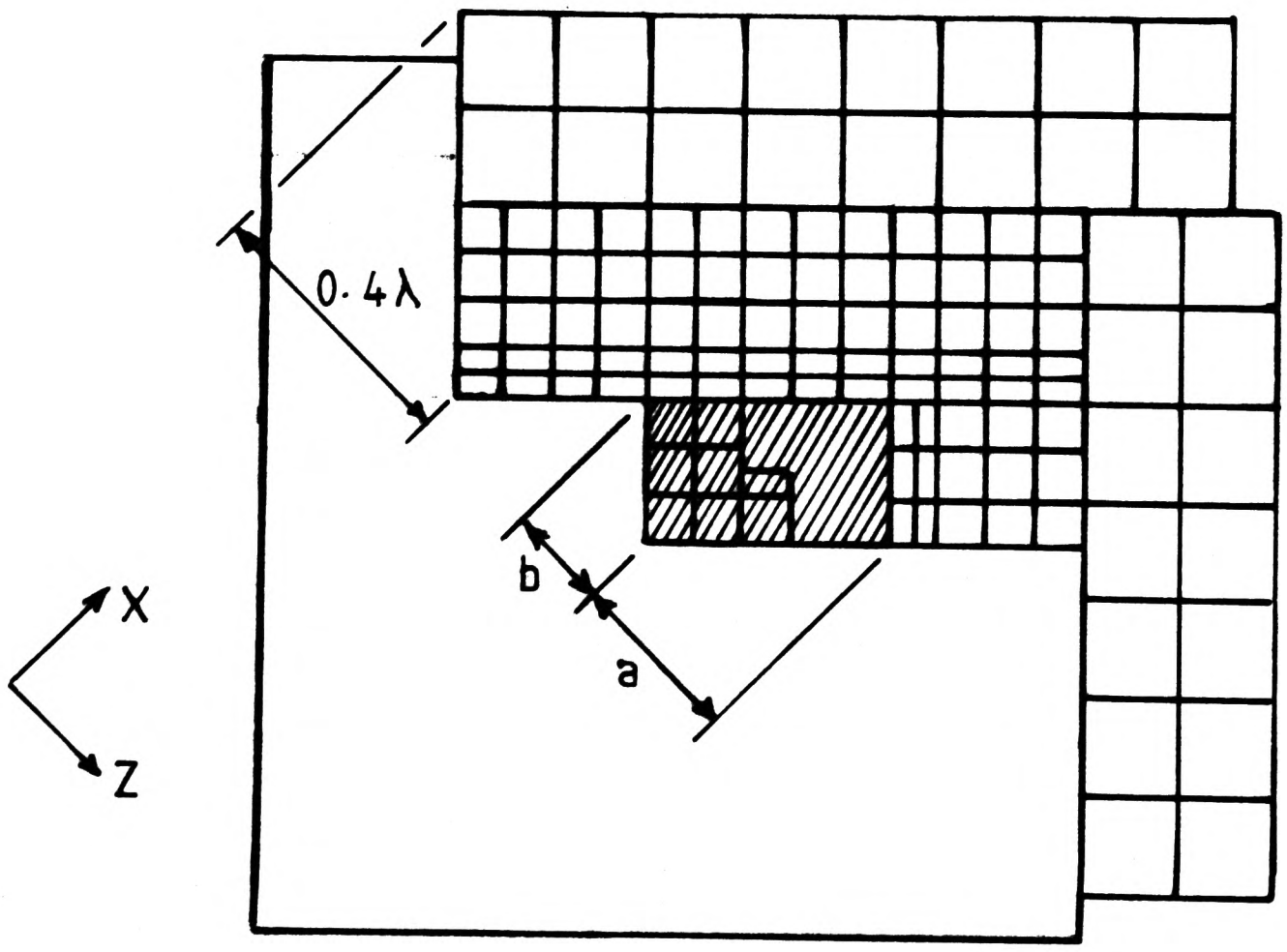


Fig. 3.12: Cells into which the hole and surrounding region are divided.

The FORTRAN program devised to find reflections from holes employed "cells" similar to the ones in Section 3, with smaller cells near the edge of the hole, as illustrated in fig. 3.12 . With the experience of the dot program, it was thought unnecessary to divide the cells into sub-cells as before, but with this exception, the same scheme to allow for distributed charge was used. Because of the large number of cells, 240 values of (**a**,**b**) took 11 hours of CPU time. $G_x \epsilon_\infty$ (equation 3.19) was found to be 0.0086 by comparision with the experimental data, in contrast with 0.0033 for dots. This will be discussed in chapter 4.

An alternative approach to find reflectivity of holes is to compute the tangential components of electric field inside the hole, and derive a relation to find the reflected wave from the electric field. This has not been attempted. The study of holes has now been passed on to another research worker.

3.7 Simplified Version of Scattering Theory

A simplified version of the model for reflections, involving analytical expressions, can be useful for understanding the mechanism of reflection. It may also be helpful for making preliminary studies, and in confirming that quantities such as change in power-flow angle are small.

The unknown potential distribution is approximated by a rectangle function as shown in fig. 3.13 . This is done simply by assuming an effective length and width larger than the physical dimensions of the dot. (The effective

dimensions of a hole are smaller than its physical size). $\mathbf{v}_d$ is taken to be the average value of $\mathbf{v}_o \exp(-jkz)$ within the effective size of the dot. The change is potential is

$$\left.\begin{aligned} \Delta V(z,x) &= V_d - V_o \exp(-jk_o z) && \text{within the dot,} \\ \Delta V(z,x) &= 0 && \text{outside it.} \end{aligned}\right\} \quad (3.28)$$

The 2-dimensional spatial Fourier transform $\Delta\bar{\mathbf{v}}(k_z, k_x)$ of $\Delta\mathbf{v}(z,x)$ is found for a particular dot size and used to find $\bar{\sigma}(k_z, k_x)$. Integrating (3.11),

$$\Delta V(z,x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi(\varepsilon_o+\varepsilon)\sqrt{(z-z')^2+(x-x')^2}}\, \sigma(z',x')\, dz'\, dx' \quad (3.29)$$

This is a convolution; taking Fourier transforms and rearranging,

$$\bar{\sigma}(k_z,k_x) = (\varepsilon_o+\varepsilon)\sqrt{k_z^2+k_x^2}\ \Delta\bar{V}(k_z,k_x) \quad (3.30)$$

where $2\pi/\sqrt{(k_z^2+k_x^2)}$ is the Fourier transform of $1/\sqrt{(z^2+x^2)}$ (Ref [6], with different notation). 90° reflectivity is found from $\bar{\sigma}(0,-k_o)$ as before.

The function used as the effective length is arbitrary, although it has certain properties which are obviously required. For example, the effective length is proportional to the length for very small lengths ($<<\lambda$), but is equal to the length plus a constant for very large dots ($>>\lambda$).

$$A = a + 0.12[1-\exp(-12a/\lambda)][1-\exp(-12b/\lambda)] \quad (3.31)$$

The parameters 0.12, 12 and 12 were selected by trial and error, to make the simplified version as close as possible to

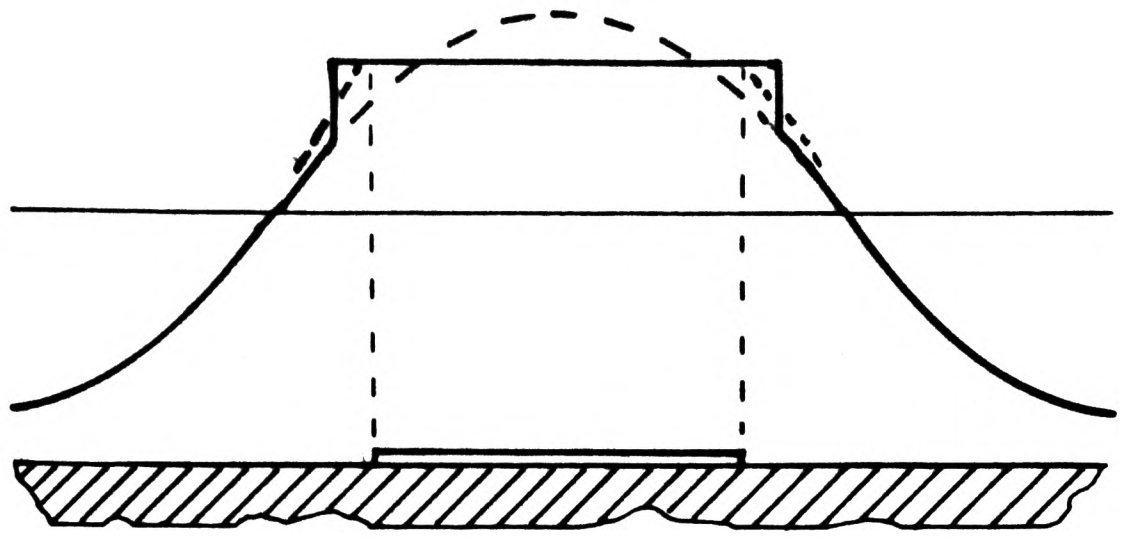


Fig. 3.13: Approximate function used for potential in the simplified version of the model.

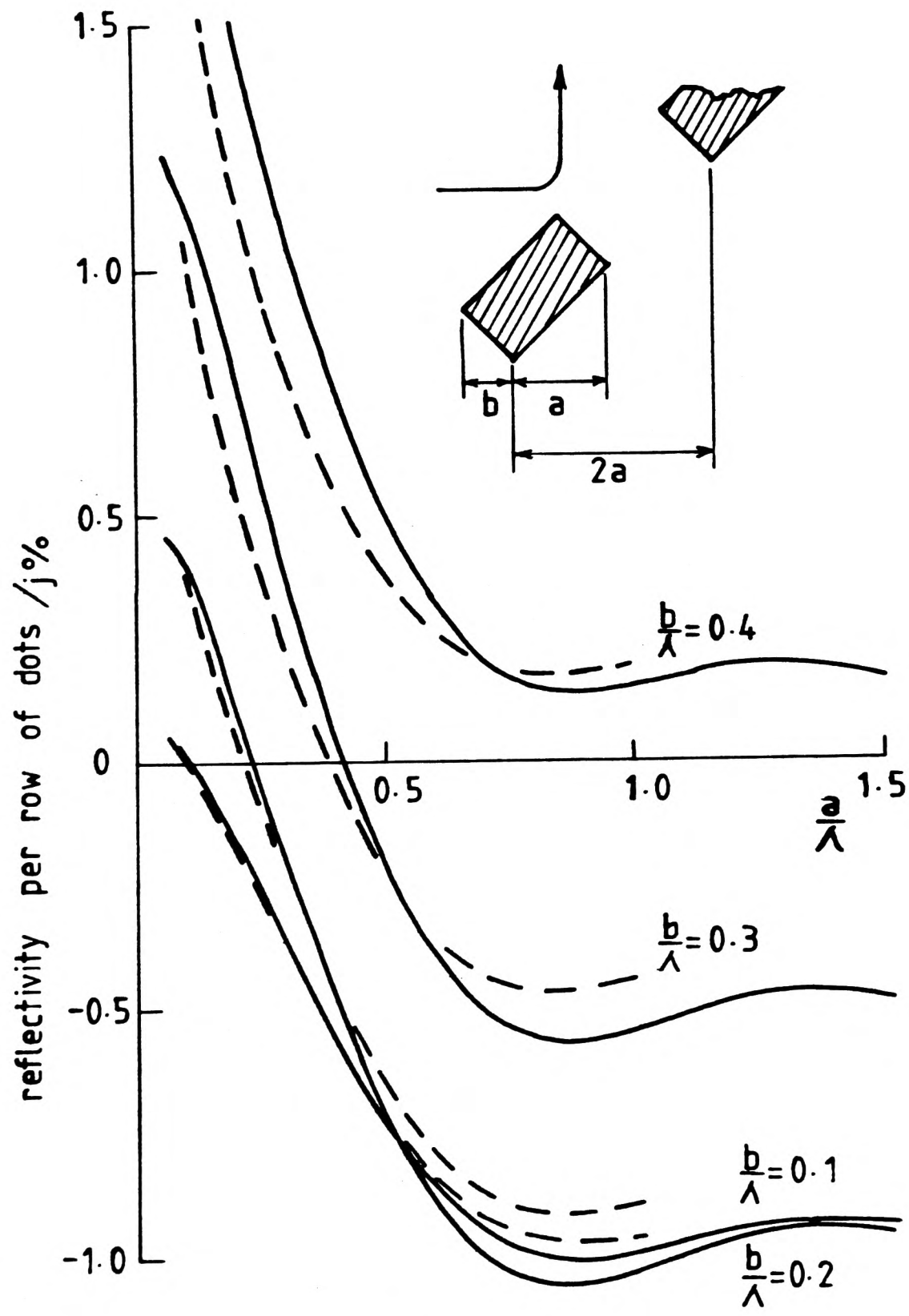


Fig. 3.14: Comparison of simplfied (___) and numerical (---) versions for dots reflecting through 90 degrees.

the numerical result for dots reflecting through 90°. The numerical and simplified results are compared in fig. 3.14 .

The expressions obtained were:

For dots,

$$jr = j\,4\pi G\,\varepsilon_\infty \frac{A\,B}{\lambda\,L_1}\left(-\mathrm{sinc}(2\beta) + \mathrm{sinc}^2\alpha\,\mathrm{sinc}^2\beta\right) \qquad (3.32)$$

For holes,

$$jr = j\,4\pi G\,\varepsilon_\infty \frac{A\,B}{\lambda\,L_1}\,\mathrm{sinc}(2\beta) \qquad (3.33)$$

where

$$\alpha = \pi A/\lambda, \quad \beta = \pi B/\lambda \qquad (3.33a)$$

and **A** and **B** are effective lengths.

It seems reasonable to assume that the same effective length can be used when caluclating scattering and change in velocity. Using the normalised value of $L=\lambda$, the reflection coefficient at angle $(\theta+45)°$ to the **Z** axis is

$$jr = j\,4\pi G\,\varepsilon_\infty \frac{A\,B}{\lambda^2}\Big(-\mathrm{sinc}[\alpha(1-C)]\,\mathrm{sinc}[\beta(1+S)] + \mathrm{sinc}\,\alpha\,\mathrm{sinc}\,\beta\,\mathrm{sinc}(\alpha C)\,\mathrm{sinc}(\beta S)\Big) \qquad (3.34)$$

where

$$C = \sqrt{2}\cos\theta, \qquad S = \sqrt{2}\sin\theta \qquad (3.35)$$

Taking a normalised value of $\sqrt{(G_z\,\epsilon_\infty G\epsilon_\infty)}=1$, this relation is compared with the numerical model for a particular dot size in fig. 3.15 , showing only approximate aggreement.

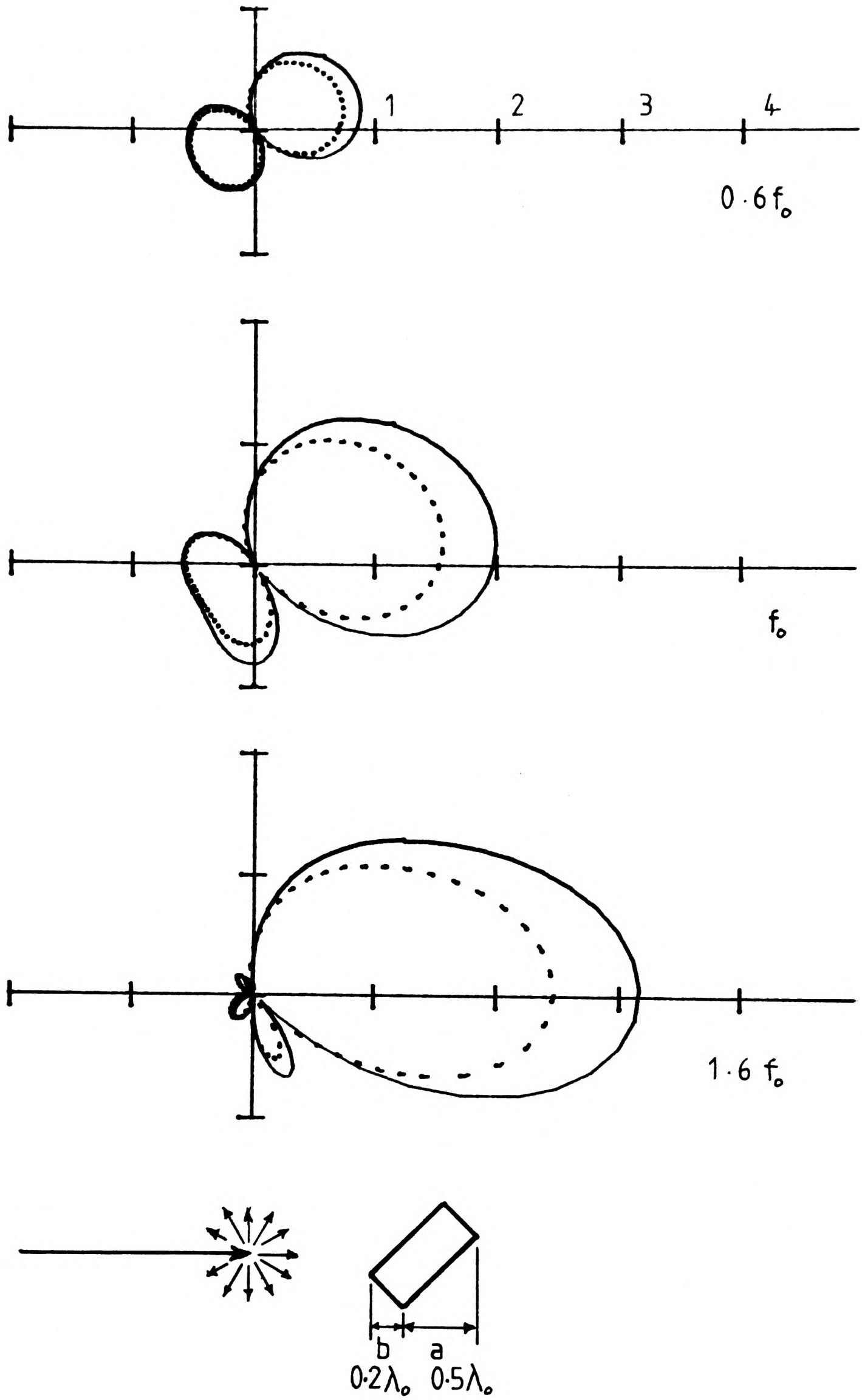


Fig. 3.15: Comparison of simplfied (____) and numerical (---) versions for scatter by a dot, showing only approximate agreement. Normalised values of $L = \lambda_0$ and $\sqrt{G_z G_\phi}\,\varepsilon_\phi = 1$ are used.

The fractional change in velocity of a wave travelling at an angle $(\theta+45)°$ to the **Z** axis in an array with n dots per square wavelength is

$$\frac{\Delta v_d}{v} = n G \varepsilon_\infty \frac{2AB}{\lambda^2}\left(1 - \mathrm{sinc}^2(\alpha C)\,\mathrm{sinc}^2(\beta S)\right) \qquad (3.36)$$

If the dots dot not have mirror symmetry about the **Z** axis, this function may not be symmetrical, resulting in a change in power-flow angle. The function can be used to confirm that this change is small.

Similar results are obtained for strips. For an isolated strip reflecting at normal incidence,

$$jr = -j G \varepsilon_\infty 2\pi \frac{B}{\lambda}\left(\mathrm{sinc}^2\beta - \mathrm{sinc}(2\beta)\right) \qquad (3.37)$$

and for a 45° inclined strip reflecting through 90°,

$$jr = j G \varepsilon_\infty 2 \sin\beta \qquad (3.38)$$

These simplified relations are not accurate for very small values of **A** and **B**. For example, (3.37) predicts that **r** is proportional to $\mathbf{B}^3$ for small values of **B**, while the accurate numerical result indicates that **r** is proportional to $\mathbf{b}^2$.

A one dimensional version of (3.30) is required in other sections of this chapter.

$$\bar{\sigma}(k_z) = |k_z|\,(\varepsilon_o+\varepsilon)\,\overline{\Delta V}(k_z) \qquad (3.39)$$

In those sections, the approximation for **V** described above is not used, so no additional assumptions are made.

3.8 Summary

A model for SAW reflections by dots is developed by modifying the theory on SAW generation in transducers. The properties of the material relevant for a SAW propagating in any given direction are specified by the SAW velocities on free and completely metallised surfaces, and where necessary, the LSW velocities under these two conditions. Dot reflectivity is calculated numerically from these values and the geometry of the dot. The model is also used to calculate scattering and velocity change. In addition, a version for holes is developed. Finally, by using an approximate value for the electric potential on the surface, simple approximate analytical relations for reflectivity and velocity change are possible as an alternative to the numerical calculation.

CHAPTER 4

RESULTS

Agreement between theoretical and experimental results for 90° reflections, scatter and velocity change due to dots in relatively sparse arrays has been good, although experimental data for scatter is limited. Previous publications also based on this project [15,16] reported the need for a correction factor of 0.7 for 90° reflections, but this has been resolved by taking the effect of the **X** propagating LSW into account. The results for 90° reflections from holes have also been encouraging, except for a scaling factor.

4.1 The Evaluation of $G\epsilon_\infty$

The theory in chapter 3 makes use of a constant $G\epsilon_\infty$. For calculating $G_z\epsilon_\infty$, the method described in Section 1 of chapter 3 is appropriate, since the coupling to the LSW is zero. Using data from ref. [1], $G_z\epsilon_\infty = \frac{\Delta v}{v} = -0.0236$.

In other directions, particularly near the **X** direction, coupling to the LSW is significant, and the modification described in section 2 of chapter 3 is required. The calculation makes use of free and metallised velocities of

both the SAW and the LSW. The result is shown in fig. 3.10 . Data from refs. [1] and [21] were used.

LSW velocities in the **X** direction were measured (chapter 2) to check the values given in [21] and [20]. These are given in table 4.1 with previously published results.

Using LSW velocities (free and metallised) obtained from this project, and the SAW velocities (v_o=3723, v_m=3691) from reference [1], $G_x\epsilon_\infty$ is computed to be 0.0042 . This is compared with the complete numerical computation [29], which does not neglect bulk wave generation, in table 4.2 . In this calculation, $\epsilon_s(s)$ is found numerically, and the relation $G_x\epsilon_\infty = \epsilon_\infty/[\ (s_o\epsilon_s'(s_o)\ ]$ [27] is used. The apostrophe denotes differentiation. Also indicated is the value obtained neglecting both bulk and LSWs (i.e. assuming $G_x\epsilon_\infty = \frac{\Delta v}{v}$).

It is clear that neglecting the LSW gives a value of $G_x\epsilon_\infty$ that is too large by a factor of about two; reflectivity, being proportional to $\sqrt{(G_x\epsilon_\infty)}$, is thus overestimated by a factor of $\sqrt{2}$. This difficulty is resolved when the LSW is taken into account.

4.2 Comparison of $G_x\epsilon_\infty$ with Dot Data

The "calculated" value of $G_x\epsilon_\infty$ found from free and metallised SAW and LSW velocities will now be compared with the experimental values obtained by a least-squares fit through the dot reflectivity data. More detailed comparisons of theory and experiment for individual dot sizes follow in later sections.

Table 4.1

LSW velocities (m/s)	v_{Lo} (free)	v_{Lm} (metallised)
Computed from material constants [20]	4670	4080
Experimental, time-of-flight [20]	4640	4170
Experimental, phase, (this project)	**4638**	**4143**

Table 4.2

Computation of $G_x\epsilon_\infty$	
Assuming $G_x\epsilon_\infty = \frac{\Delta v}{v}$, $\frac{\Delta v}{v}$ from [1]	- 0.0086
From experimental values of v_o, v_m, v_{Lo}, v_{Lm}	**- 0.0042**
Using $G_x\epsilon_\infty = \epsilon_\infty / [\ s_o e_s'(s_o)\ ]$, numerical [29]	- 0.0037
$G_x\epsilon_\infty$ by Least Squares Fit through Dot Data	
Reflectivity, figs. 4.1,2,3,5	- 0.0036
Reflectivity, fig. 4.8	- 0.0031
Reflectivity, both sets	**- 0.0033**
X-velocity, fig. 4.11	- 0.0033

90° reflectivity is proportional to $\sqrt{(G_X \epsilon_\infty)}$ and X propagating SAW velocity to $G_X \epsilon_\infty$. The least squares fit through the data suggest values of $G_X \epsilon_\infty$ given in table 4.2 . ($G_Z \epsilon_\infty$ is assumed to be correctly determined by the theory. This assumption is supported by data on Z velocity slowing by dots.)

In this context, the "calculated" value and the reflectivity data agree to within a factor of $\sqrt{(0.0033/0.0042)} = 0.89$. In the sections that follow, individual readings of dot reflectivity will be compared with the theory multiplied by 0.89 . For similar reasons, measurements of X velocity change due to dots will be compared with theory multiplied by 0.0033/0.0042 = 0.79 .

It should be noted that, with regard to the two sets of reflectivity data in table 4.2, although the scatter of the first set was much larger than the second, the number of readings is also much larger. The random errors of the two values of $G_X \epsilon_\infty$ are approximately equal.

4.3 Reflections through 90° and Scattering by Widely-Separated Dots

Experimental and theoretical results for widely-separated dots reflecting through 90° appear in figs. 4.1 and 4.2 . Further results, from the pilot device, do not differ significantly, but are recorded in table 4.3 for completeness. Only the amplitude of reflectivity was verified experimentally; the polarity is inferred from the theory.

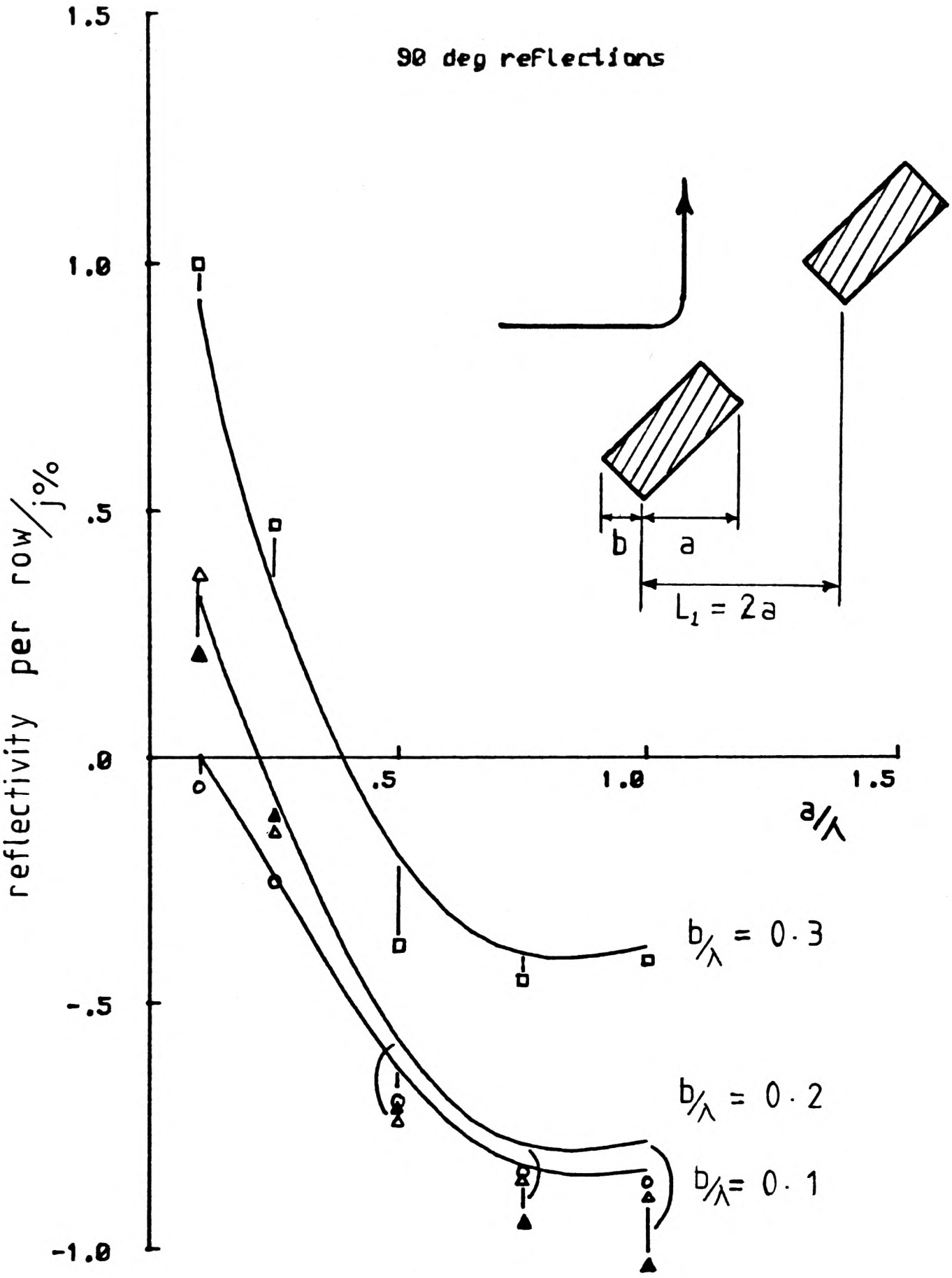


Fig. 4.1: 90 degree reflectivity of widely-separated dots.
Points: Experiment
Lines : Theory x 0.89
Error in Dot size: (▲), 0.03 wavelengths
others, 0.01 wavelengths

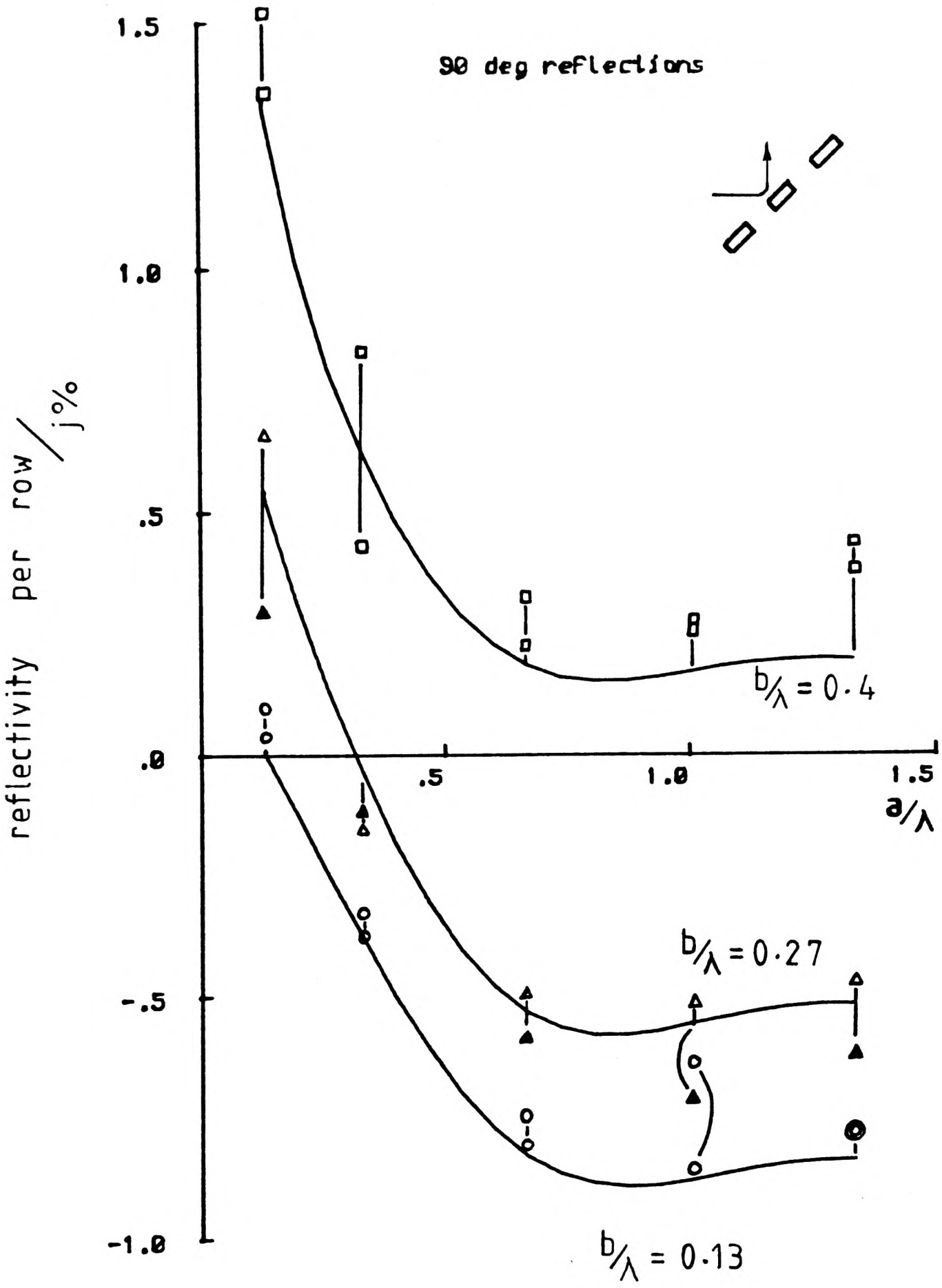


Fig. 4.2: 90 degree reflectivity of widely-separated dots. (See caption, fig. 4.1)

Table 4.3: Results of Pilot Device
Reflectivity through 90° /j%

a and b in wavelengths

b	a	theory ×0.89	expt	b	a	theory ×0.89	expt
0.21	0.21	0.06	0.06	0.21	0.5	-0.54	-0.63
0.25	0.25	0.11	0.09	0.25	0.6	-0.54	-0.61
0.29	0.29	0.18	0.12	0.29	0.7	-0.44	-0.49
0.33	0.33	0.28	0.16	0.33	0.8	-0.24	-0.30

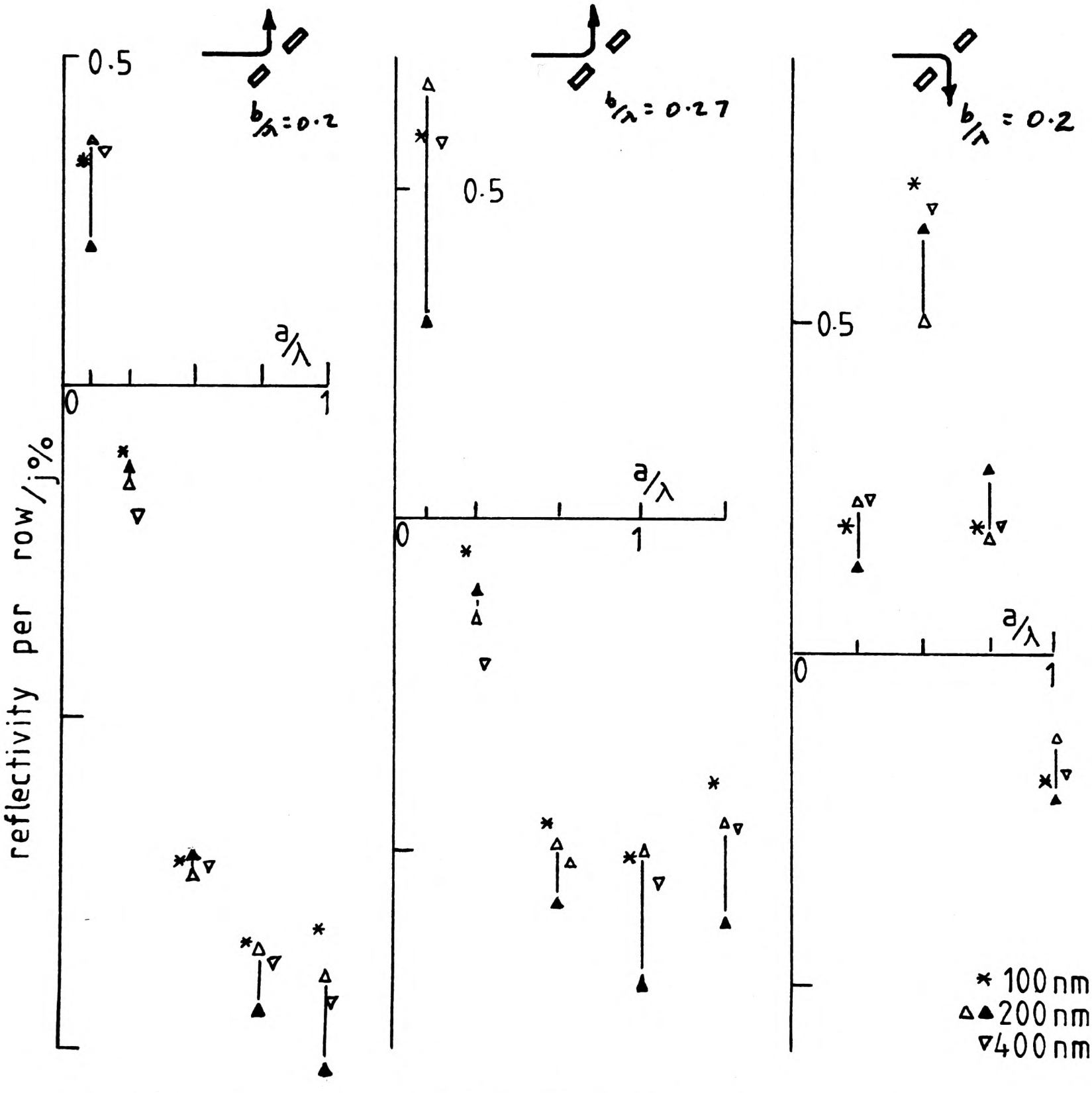


Fig. 4.3: Variation of reflectvity with dot thickness, through +90 and -90 deg. (See caption, fig. 4.1)

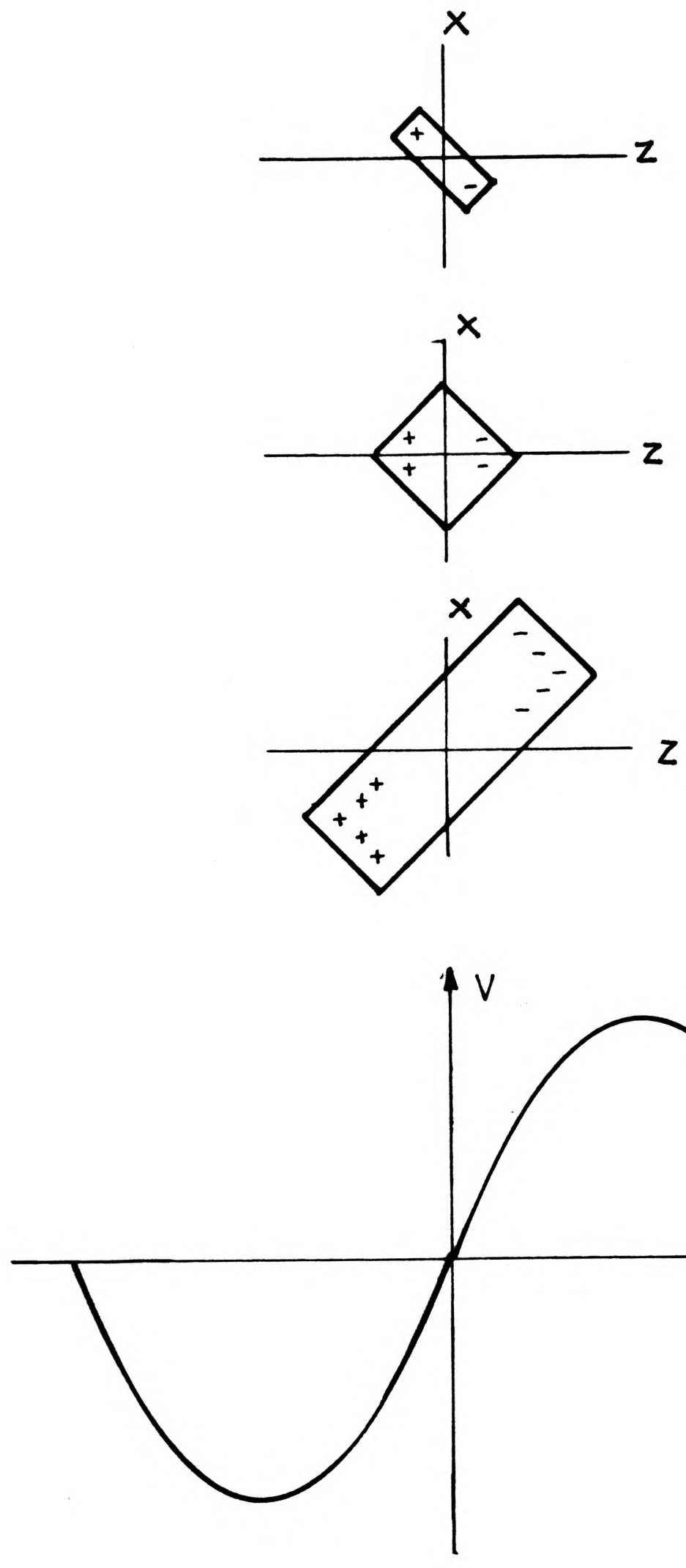


Fig. 4.4: Charge distribution due to potential of incident wave, showing little movement of charge in the X direction when the dot is square.

Since the rows of dots were widely separated in these experiments (Section 1, Chapter 2), interactions between rows were obviously negligible. Capacitative interaction between dots in the same row was calculated to be small. Varying the number of dots per row experimentally as described (Chapter 2) also failed to reveal any variations comparable to the scatter in the results (after appropriate scaling). The effect of mass loading is also known to be small, as deliberate vaiations in metal thickness between 100nm and 400nm did not result in significant variations in reflectivity. (See fig. 4.3 .) Error in dot size is less than about 0.01 wavelength, except where noted.

The reason for the scatter in experimental points is not known. It is not due to transducer characteristics, which cancel out in the calculation, nor is it caused by multiple reflections, which should be similar in nominally identical devices.

The results confirm the large range of reflectivities expected. A prominent feature is that the reflectivity of small dots is particularly low when the dots are square. This is illustrated in fig. 4.4 . Equation (3.19) implies that an **X**-propagating wave is generated only by charge redistributed in the **X** direction; this is small in the case of a square dot.

Useful dot sizes can now be identified. If the computed data is re-plotted as curves with constant (**a**-**b**), the curves show stationary points approximately when **b**=0.2λ. If **a** dot array device is over-etched, **a** and **b** are both reduced such

that (**a**-**b**) is approximately constant. The reflectivity of a dot with **b**=0.2**λ** would therefore be relatively insensitive to such overetching.

If mass loading rather than electrical shorting were chosen as the reflection mechanism, it would be desirable to keep electrical reflections low so that only one mechanism would have to be dealt with. Square dots would be ideal. This shape appears to have been fortuitously chosen by Solie [36-39,44] in some successful devices.

If a device is to be weighted by the number of dots per row, it may be necessary to have a low reflectivity per dot to reduce quantisation errors. A size of (for example) a=0.9**λ**, b=0.3**λ** would have low reflectivity and also low scattering (see section 4). With this particular weighting scheme, variations of reflectivity with overetching may be unimportant since all reflecting rows are affected in the same proportion.

Results for reflections through -90° appear in fig. 4.5 . Since the relations used in the calculations are similar to 90° reflections, the results offer further evidence that the theory for reflections is valid.

Experimental data on scattering is limited to 90°, -90° (above) 0° (velocity change, section 5), and 180°. The results for 180° reflections are given in fig. 4.6 (**Z**-**Z**) and 4.7 (**X**-**X**).

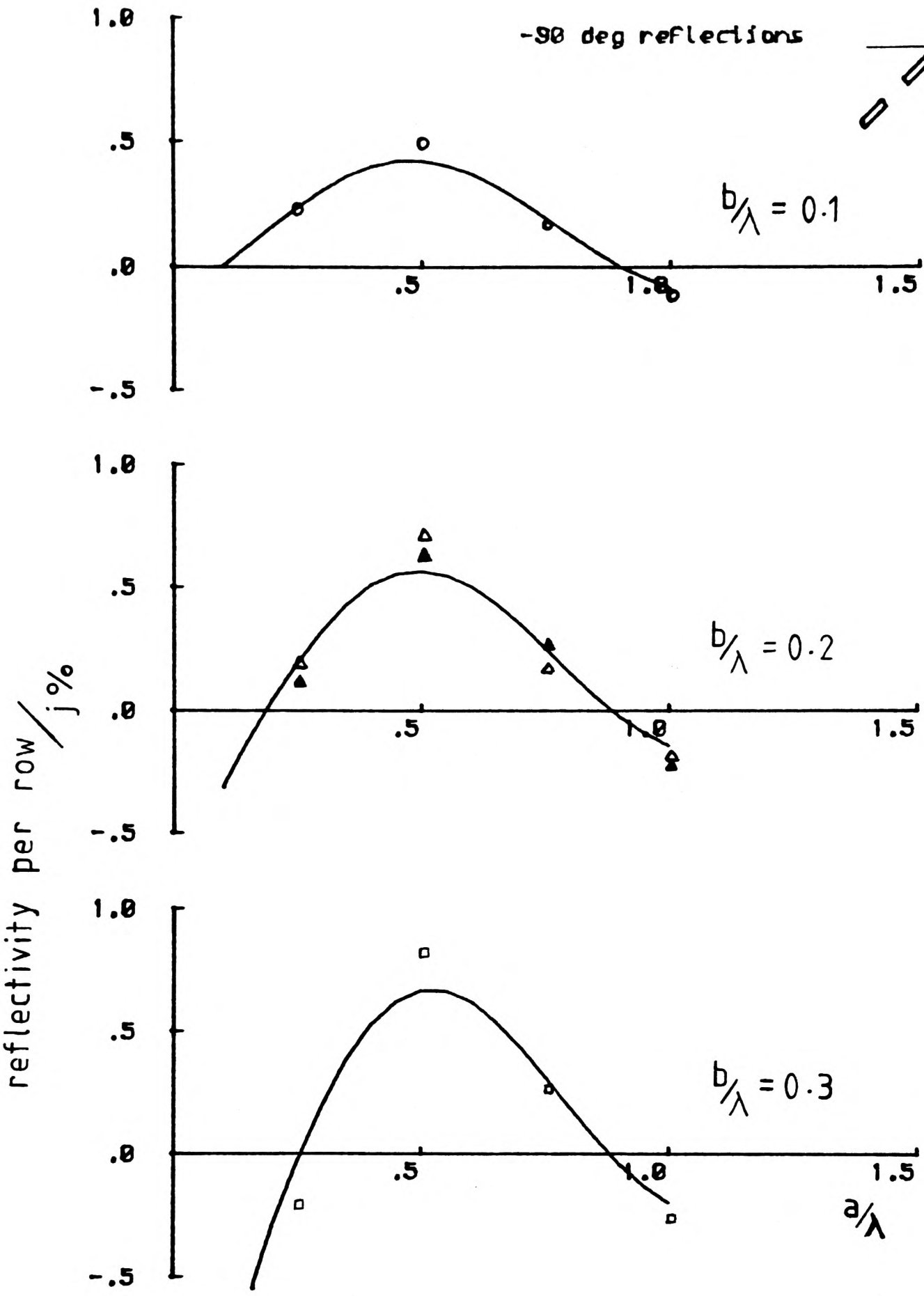


Fig. 4.5: -90 degree reflectivity of widely-separated dots.
(See caption, fig. 4.1)

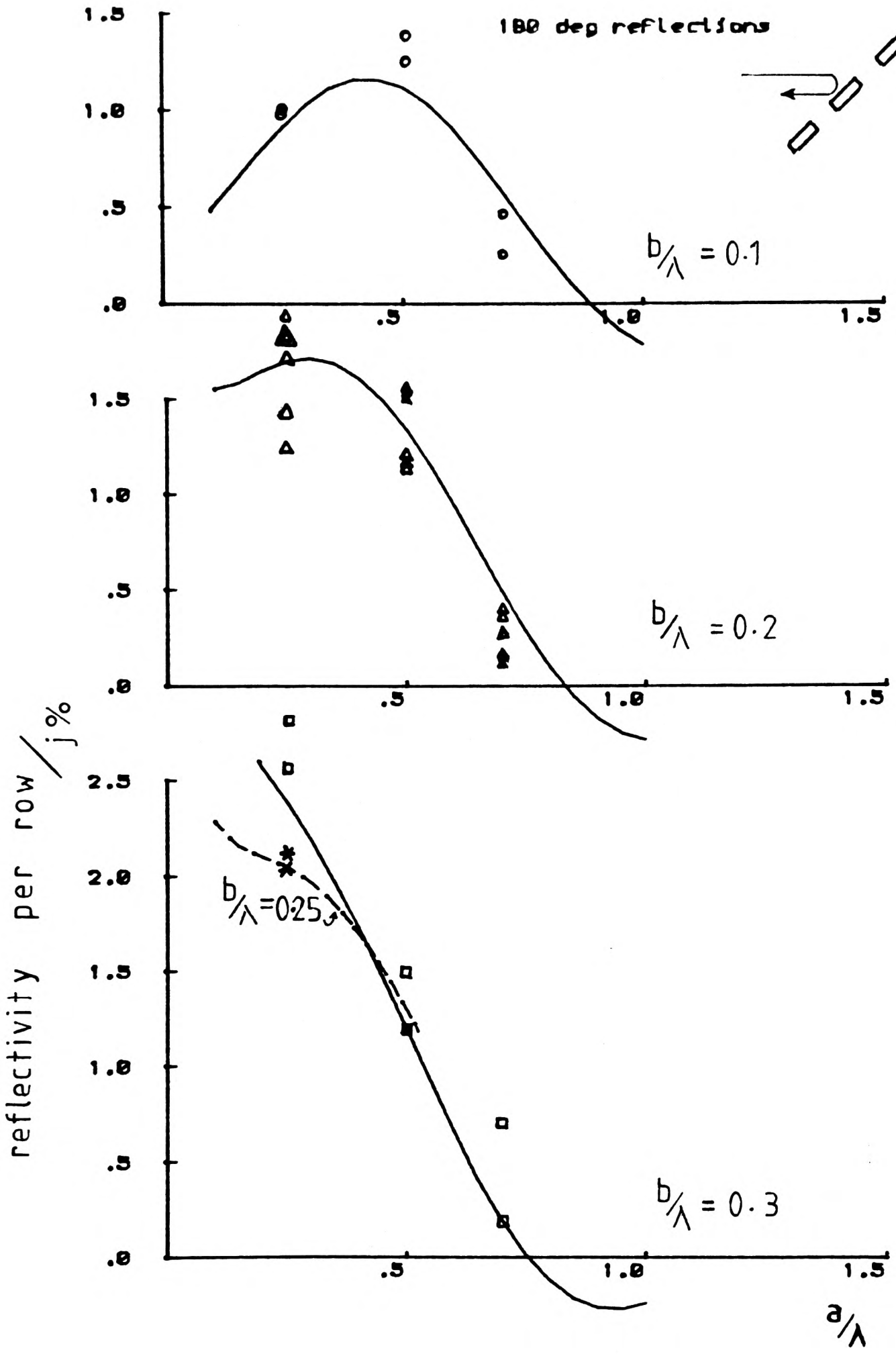


Fig. 4.6: 180 degree reflectivity of dots (Z-Z).

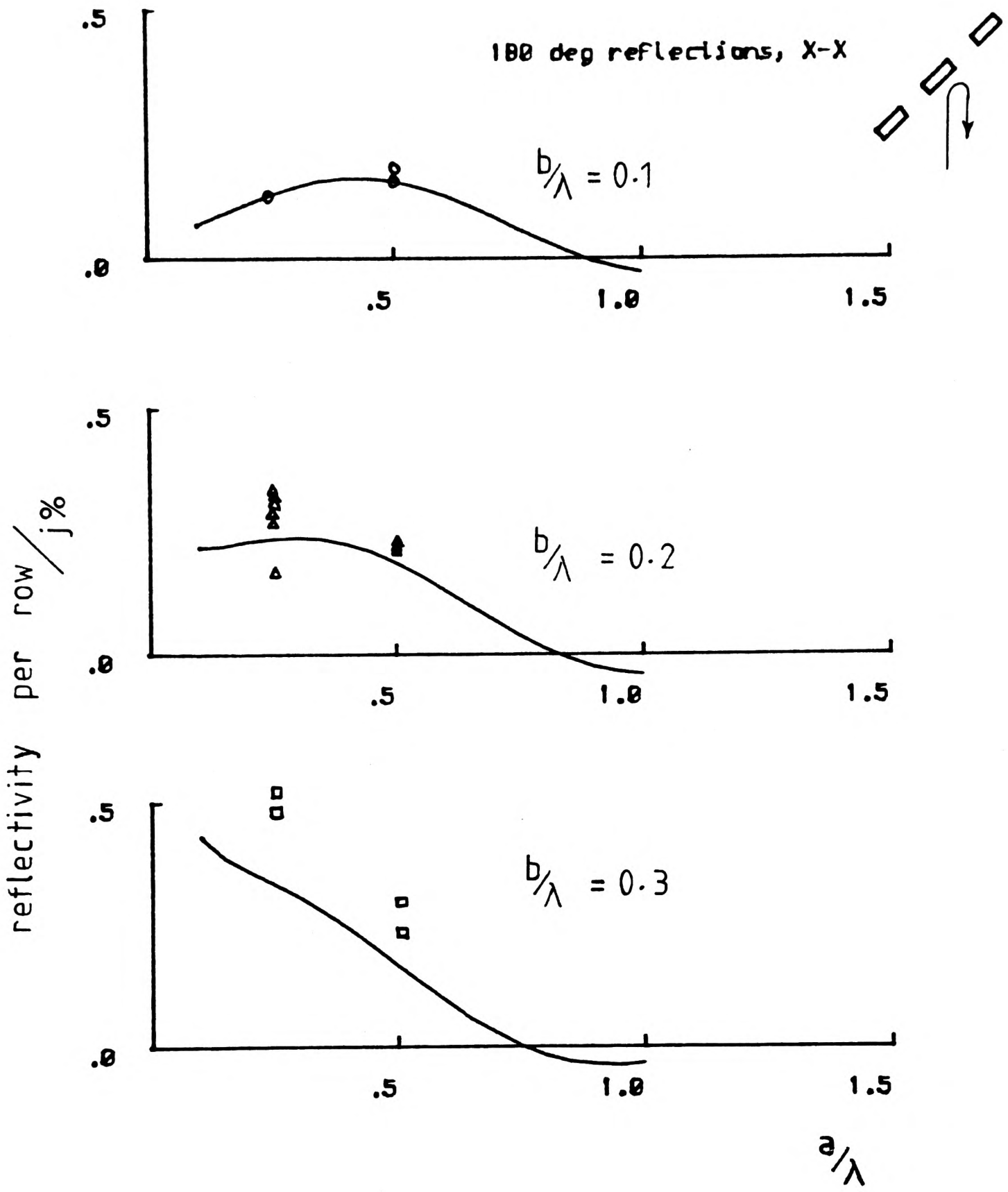


Fig. 4.7: 180 degree reflectivity of dots (X-X).
Theory x 0.79 and experiment.

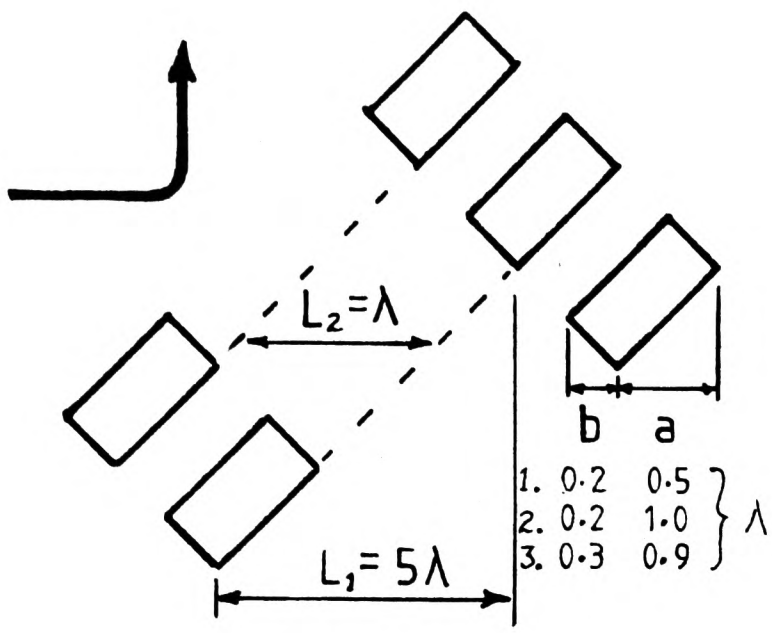
4.4 Additional Results on Reflectivity, Including Dots with Interactions

Fig. 4.8 presents the results of the second, more accurate set of experiments on reflectivity. The data is consistent with the assumption of capacitative interactions between dots. No evidence of other interaction mechanisms is apparent.

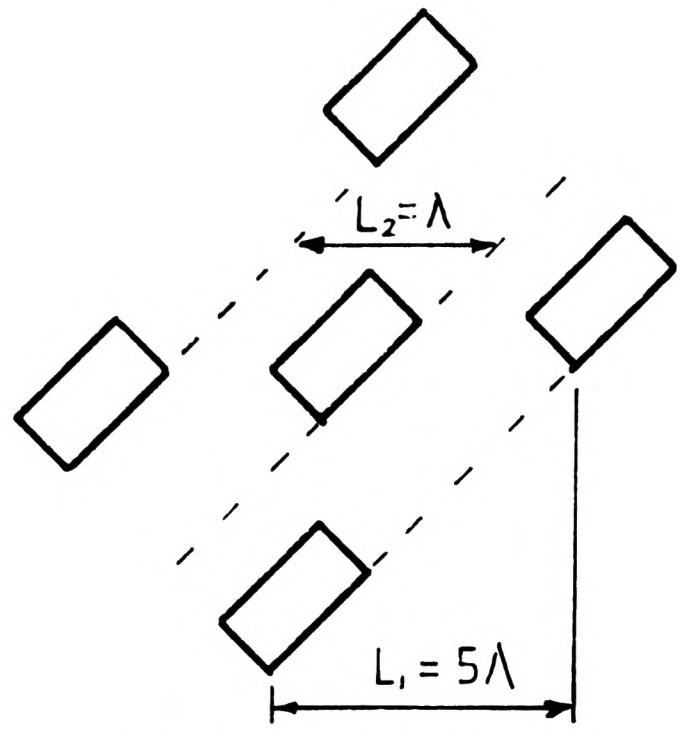
An example of a theorectical polar diagram for scatter by a dot with **a**=1.0λ, **b**=0.2λ and nearest neighbours at (0.29λ,-0.71λ), (0.57λ,-1.43λ) and inversion symmetry points is shown in fig. 4.9 . It shows that scattering is relatively low except in the first quadrant. This is useful since scattering is undesirable. Most long dots (say a>0.7λ) have this property.

4.5 Velocity Change Due to Dots

Data for velocity perturbation in the **Z** direction is presented in fig. 4.10 . Dispersion is present and would have to be taken into account in the design of a wideband device. It can be seen that the variation of velocity with frequency is particularly rapid for small dots (fig. 4.10(i)). The perturbation of potential by a small dot is slight because if the dot does not span most of a wavelength, its potential floats with the wave. As frequency increases, wavelength decreases and the dot spans a larger fraction of a wavelength, and the perturbation becomes larger.



Points marked (x)



Points marked (o)

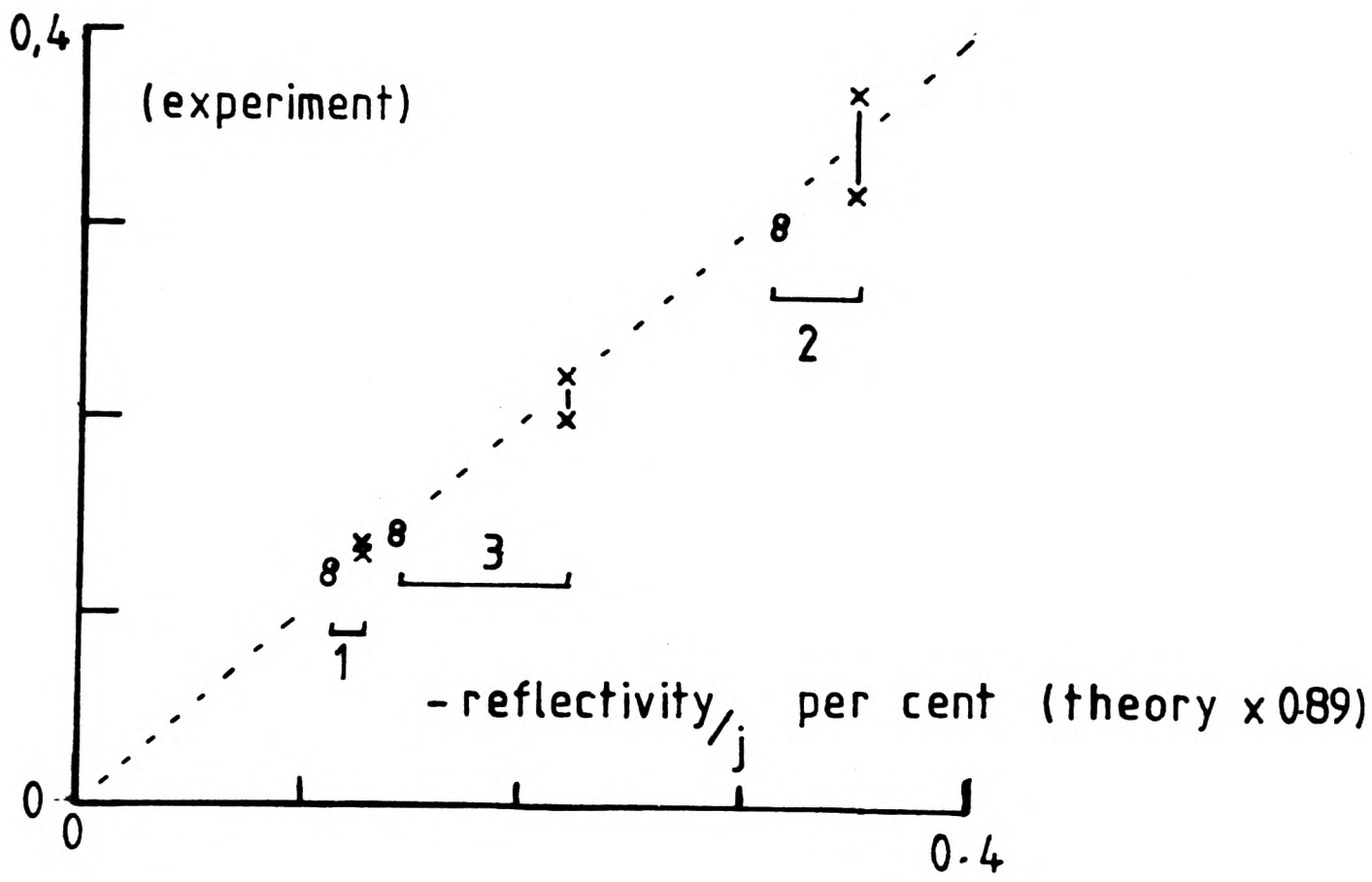
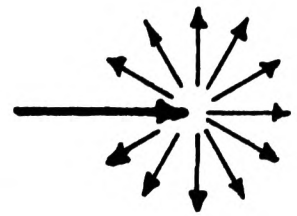


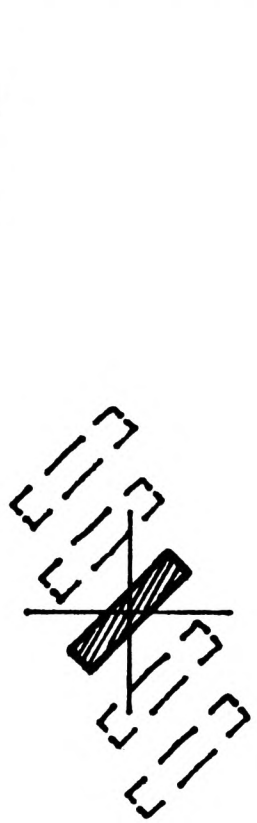
Fig. 4.8: 90 degree reflectivity, additional experiments.

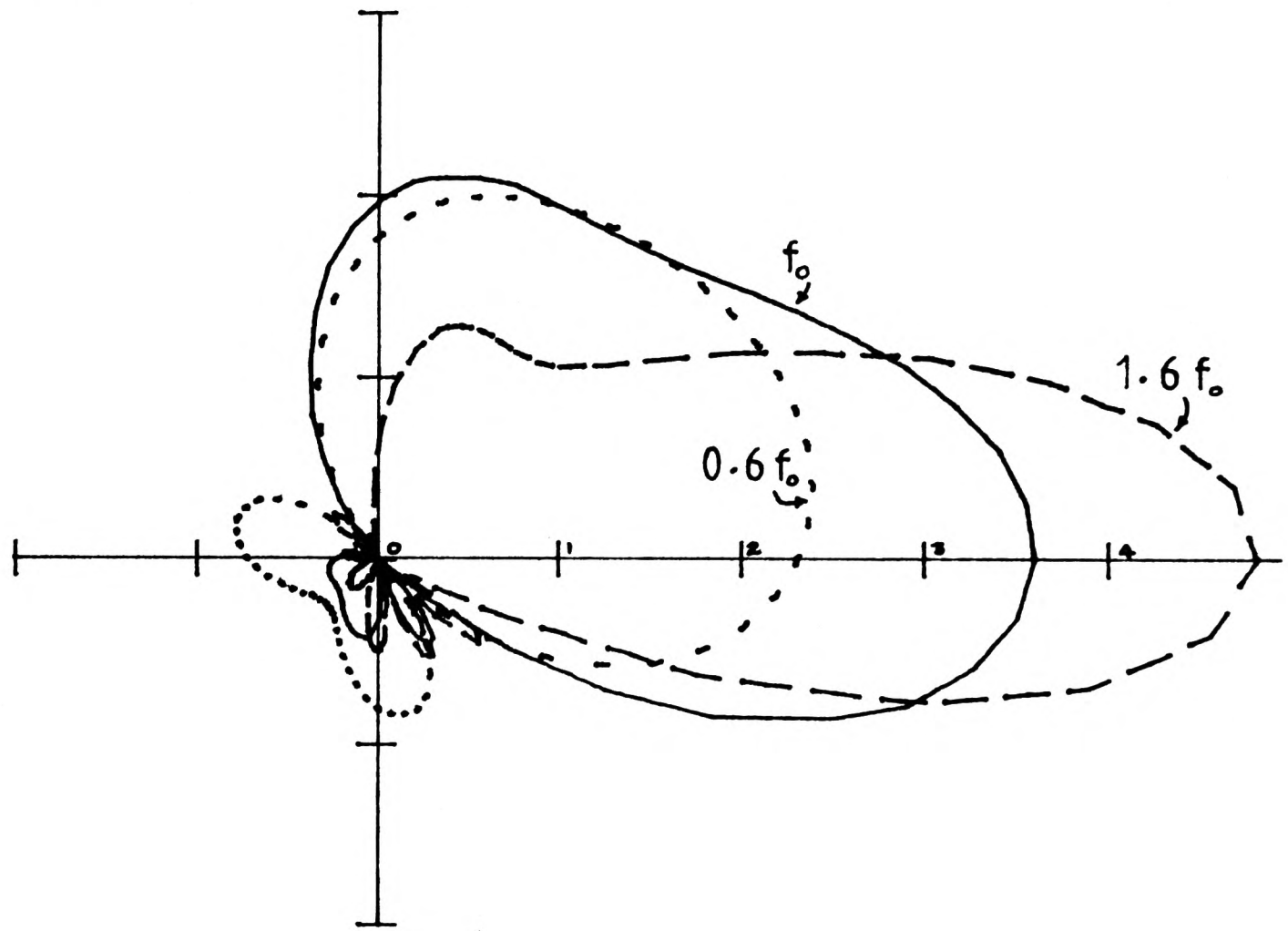
Fig. 4.9: Normalised scattering from a dot



a= 1.00λ_o
b= 0.20λ_o
Nearest neighbours:
(-0.29λ_o, 0.71λ_o)
(-0.57λ_o, 1.43λ_o)
and inversion-symmetry points
L=λ_o

$\sqrt{G_z G}\ \varepsilon_\infty = 1$





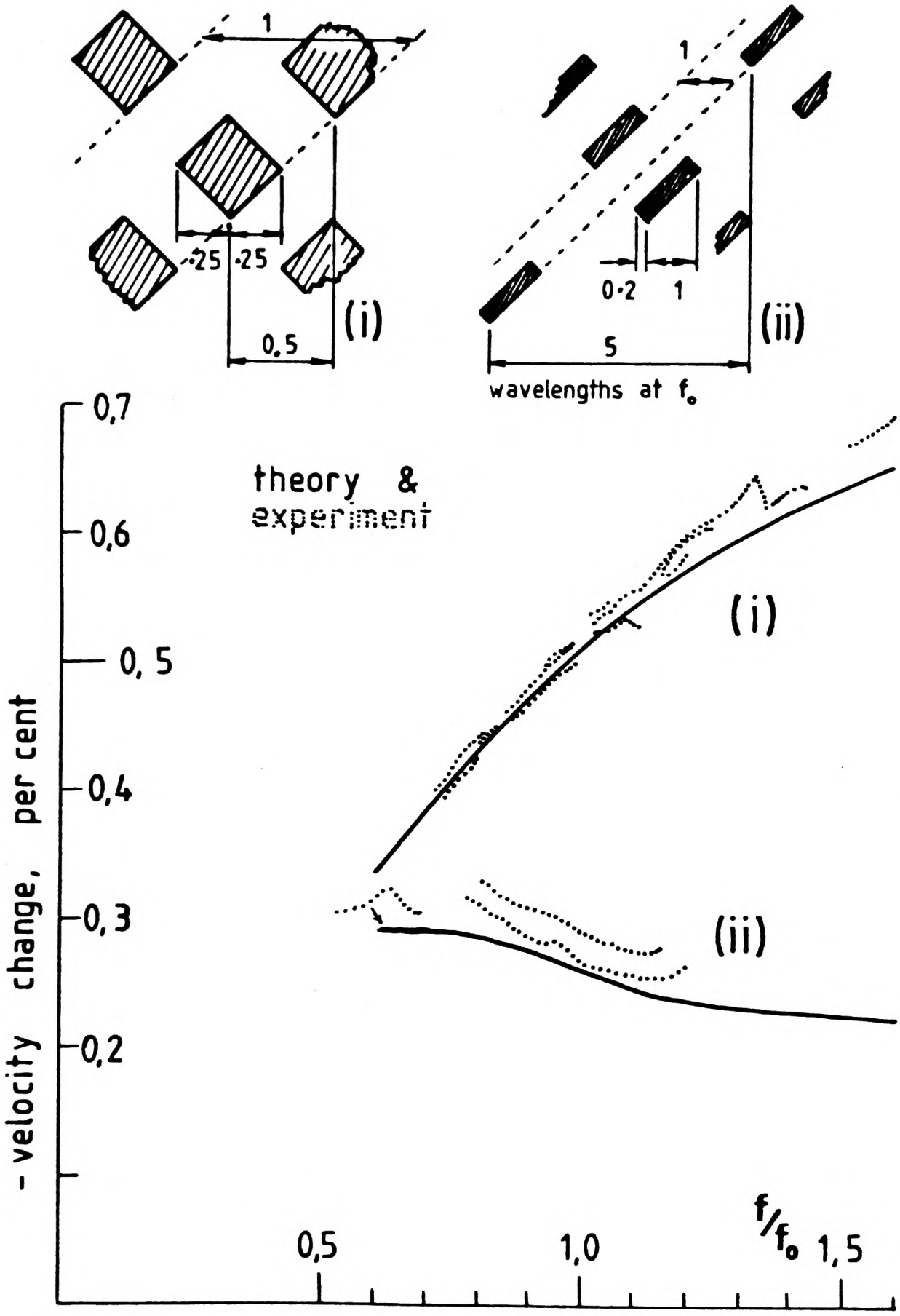


Fig. 4.10: Z-Velocity change due to dots.

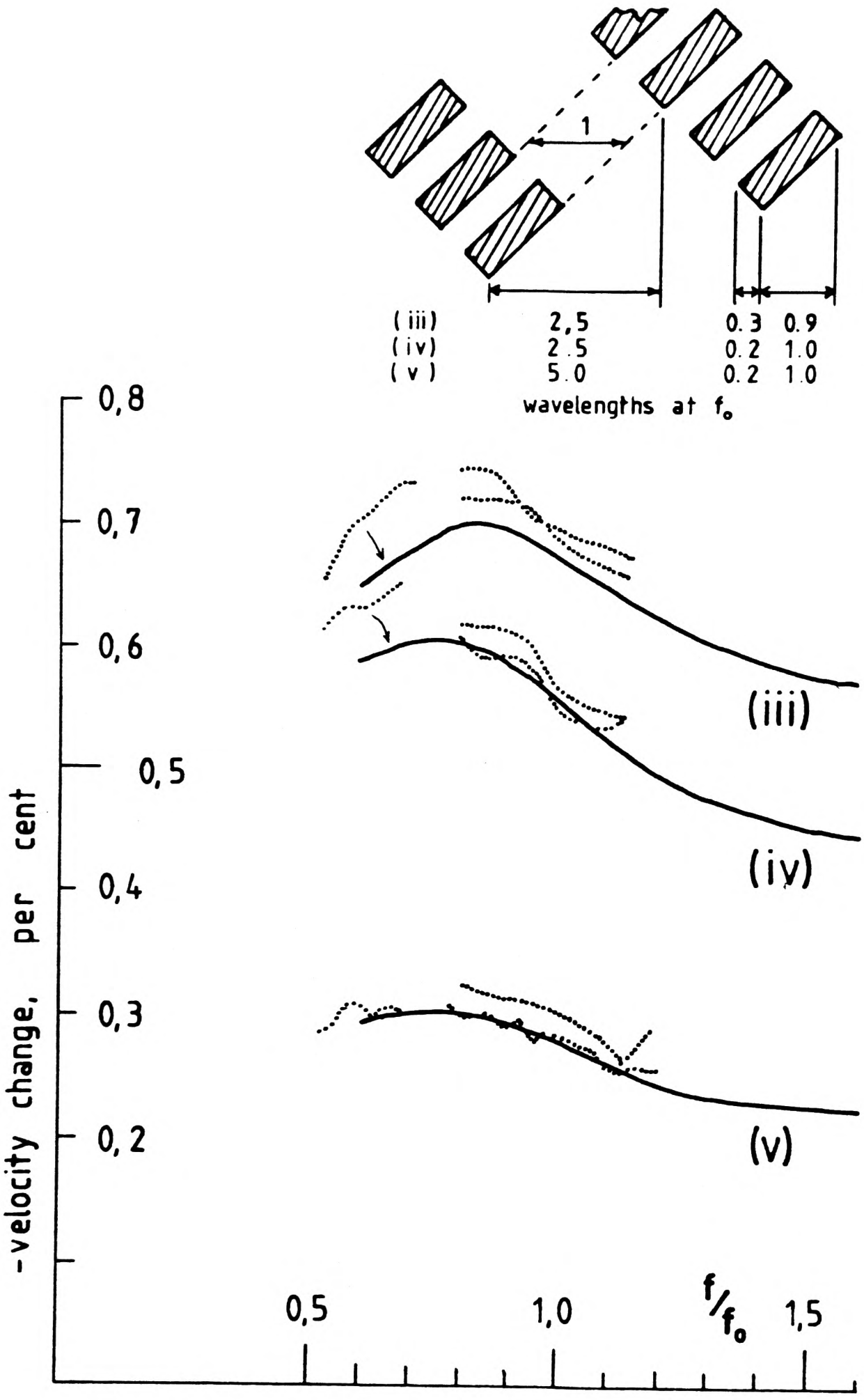


Fig. 4.10 (continued): Z-Velocity change due to dots.

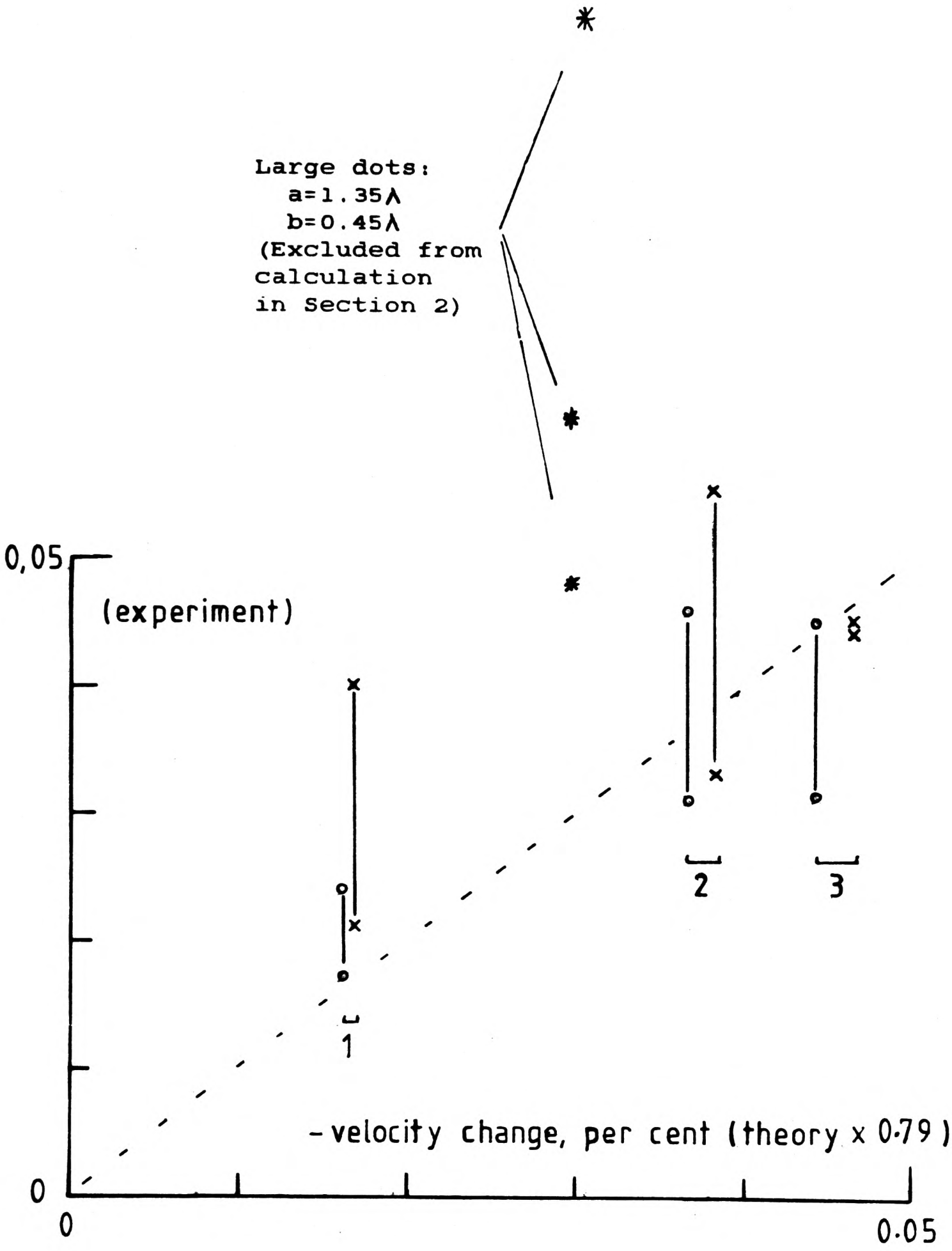


Fig. 4.11: X-Velocity change due to dots.
Dot sizes: see fig. 4.8

It would be useful to have a shape of dot such that the velocity perturbation is insensitive to overetching, but no such shape has yet been discovered.

Results of velocity change in the **X**-direction are recorded in fig. 4.11 . Because of the smaller coupling in the **X** direction, it was not thought necessary to find the variation of velocity with frequency.

There is evidence that the theory is not valid for large dots. The three experimental points for a dot size of **a**=1.35λ, **b**=0.45λ, in fig. 4.11 suggest that velocity change is larger than calculated using $G_x\epsilon_\infty$ =0.0033 . This is not surprising. The velocity change in the **X** direction due to complete metalisation is also larger than that given by equation (3.19), with this value of $G_x\epsilon_\infty$. The inference is that velocity change is proportional to $\bar{\sigma}(-k_o,0)$ in the **Z** direction, but approximately proportinal to $\bar{\sigma}(0,-k_o)$ in the **X** direction only for dots within a certain size limit. Most of the sizes represented in fig. 4.11 appear to be within this limit.

4.6 Reflections by Holes

Holes are not the exact opposite of dots for two reasons. The potential on a small dot "floats" with the incident wave, while there is no analogous effect for holes. The influence of the metal on potential extends to just outside a dot and just inside a hole, so that the effective size of a dot is larger, and the effective size of a hole smaller than their physical dimensions.

Data on hole reflectivity are given in fig. 4.12, and differ from dot reflectivity as expected. In addition, the effective value of $G_x\epsilon_\infty$ is found by comparison with the experimental data to be

$$G_x\epsilon_\infty \simeq -0.0086 = \frac{\Delta v}{V_x} \qquad (4.1)$$

assuming $G_z\epsilon_\infty = -0.0236$. Since, for dots, $G_x\epsilon_\infty < (\frac{\Delta v}{v})_x$, it might be expected that the opposite, $G_x\epsilon_\infty > (\frac{\Delta v}{v})_x$ applies fo holes. Initial results from experiments on velocity change from holes [41] support this argument. The reason for the above value of $G_x\epsilon_\infty$ for reflections is not known.

The computer program for holes also calculates

$$\sum \Delta q \Big/ \sum |\Delta q_h|$$

where Δq is the charge perturbation in any "cell" (as described in Chapter 3) and Δq_h is the perturbation in a cell within the hole. This quantity is a measure of the total charge perturbation due to the hole. The value was less than 0.06 in all cases. This indicates that the total charge perturbation due to a hole is zero, the small values being attributable to quantisation errors in the computation.

Computed results for 180° reflectivity and velocity change for holes are given in figs. 4.13 and 4.14 respectively. No experiments were performed.

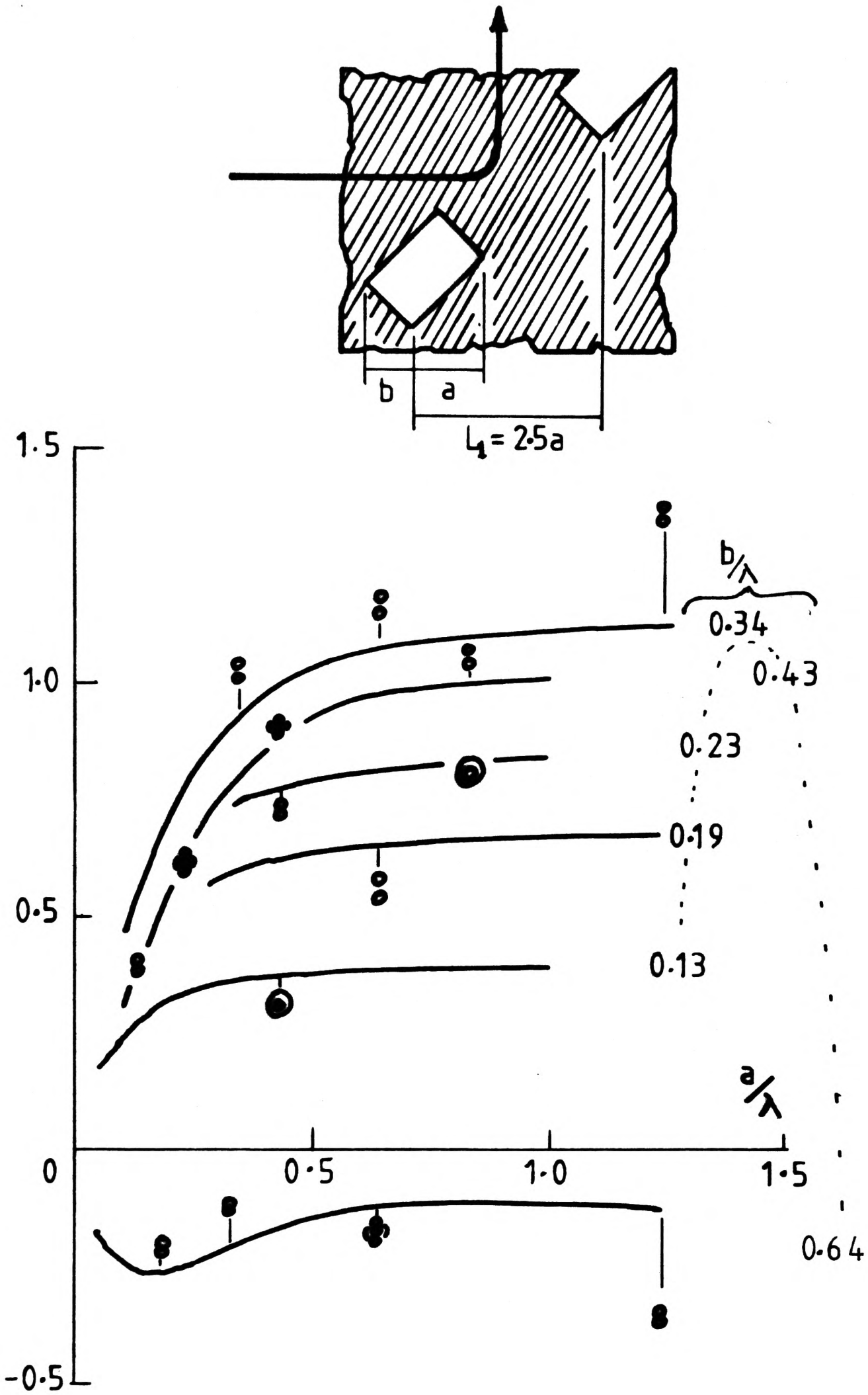


Fig. 4.12: 90 degree reflections from holes.
Points: experiment
Lines : theory (scaled)

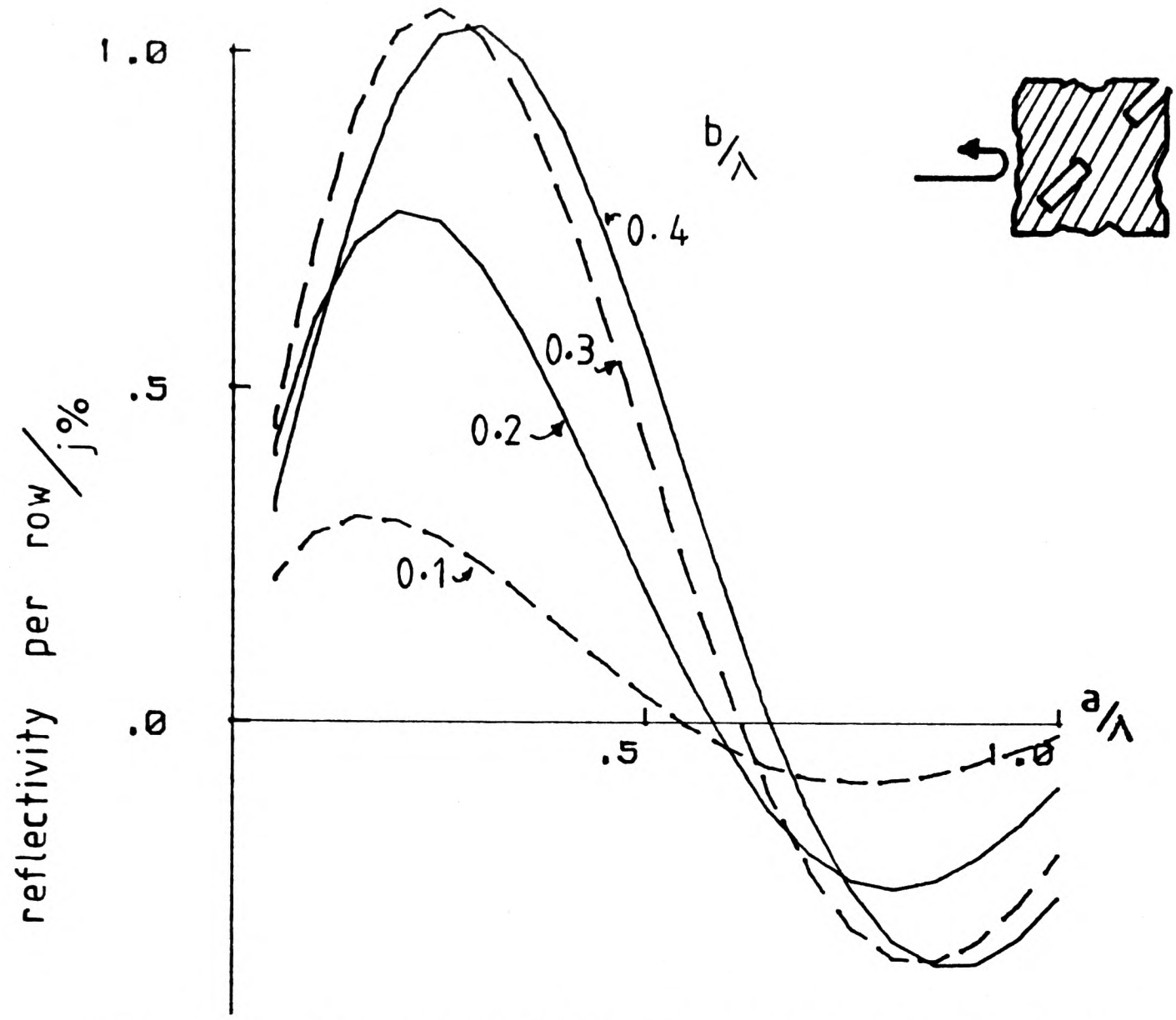


Fig. 4.13: 180 degree reflections from holes (Z-Z)
(theorectical only). Dot dimensions as fig. 4.12

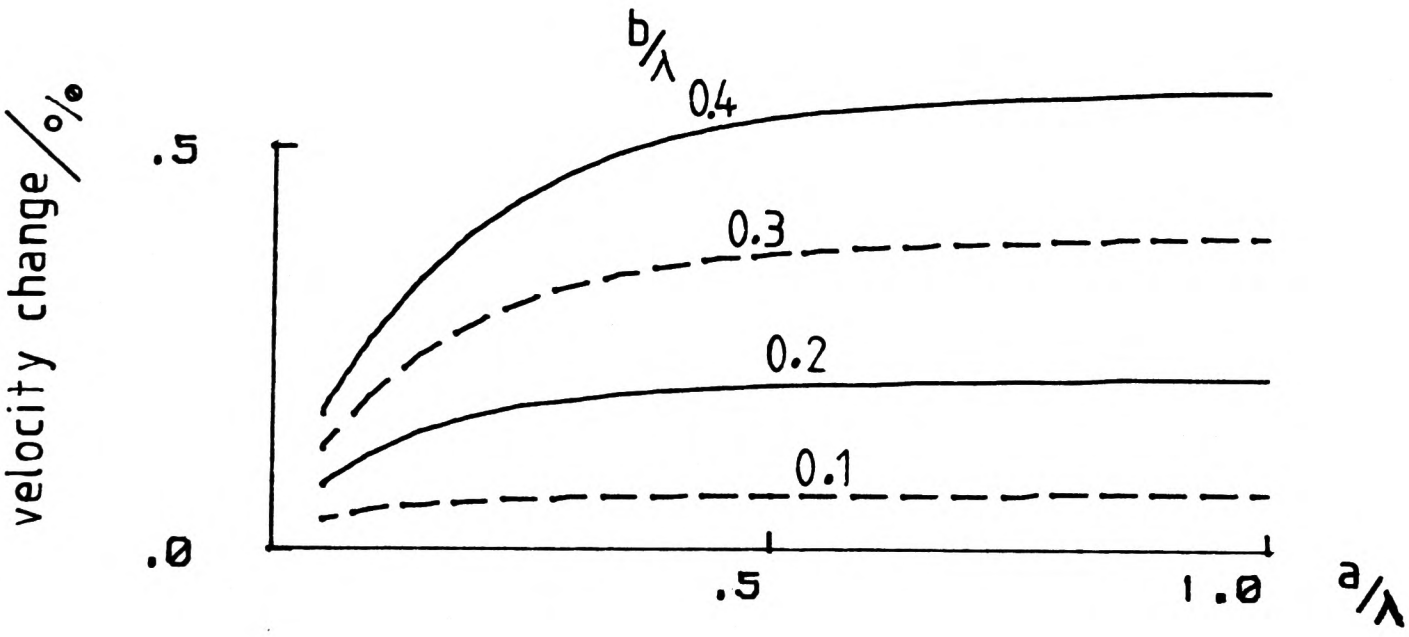


Fig. 4.14: Velocity change due to holes (Z-Z)
(theorectical only). Dot dimensions as fig. 4.12
Density of rows is 1 row per wavelength.

Because of the larger value of $G_x\epsilon_\infty$, larger reflectivities are attainable with holes than dots. This enables smaller holes to be used, probably reducing the sensitivity of **Z** velocity to overetching. The curves for 180° reflections (with **b** constant, fig. 4.13) come to zero at a smaller value of **a** than with dots. This might also be an advantage since scattering is undesirable.

One further consequence of the results for holes will be mentioned. It has already been indicated that the value $G_x\epsilon_\infty$=0.0033 may not be applicable for large dots. The data on holes confirms that $G_x\epsilon_\infty$ for dots is not valid when the fraction of metallised surface is large.

4.7 Plots of Transducer Radiation Resistance

The radiation resistance as a function of frequency is plotted for various propagation directions is fig. 4.15 . The measurements were carried out on transducers of varying aperture and period, but the results have been normalised to take this into account. The ripple on the curves for directions 74° and 16° to **Z** are due to reflections of bulk waves from the bottom of the crystal. They appear on these two graphs because the substrate used was 0.5mm thick instead of the usual 2mm.

A qualitative idea of the energy converted to SAWs, LSWs and Bulk modes in various directions can be obtained from this data. The lowest-frequency peaks correspond to the generation of SAWs; coupling is strongest in the **Z** direction, as expected. The peak due to the LSW occurs at a frequency about

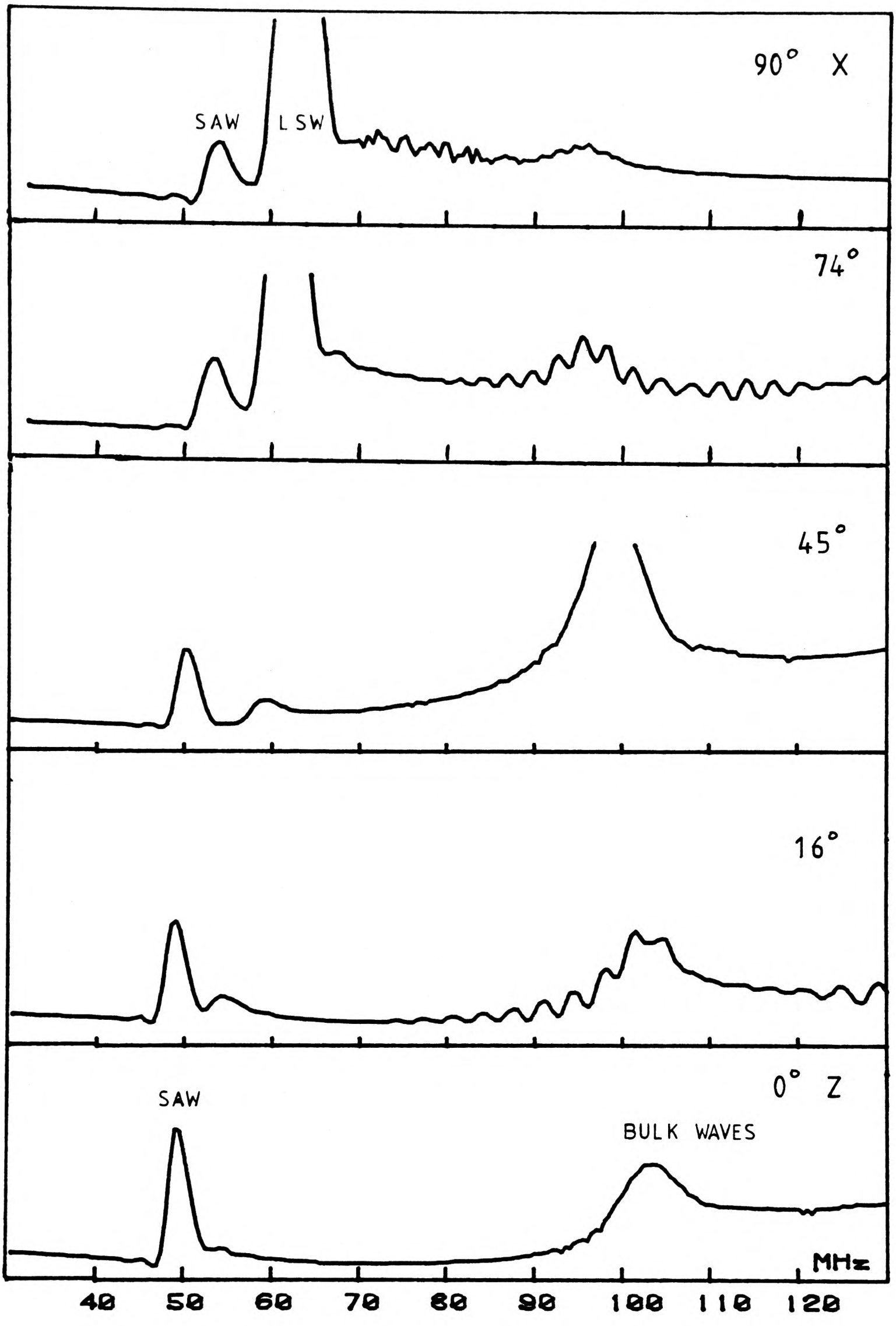


Fig. 4.15: Radiation resistance of transducers. Frequency scale is normalised to a transducer period of 70 microns.

10% higher than the SAW. It is insignificant near the **Z** direction, but very much larger than the SAW peak near **X**. Conversion to Bulk waves occurs over a wide band of frequencies, but is particularly large at about twice the frequency of the SAW peak for propagation directions near **Z**.

4.8 Reflections by Strips through 180°

When developing the model for dot reflections, it was convenient to begin with the simpler configuration of strips reflecting at normal incidence, in which charge varies in only one dimension. The principles involved were used earlier by Morgan [30] in the analysis of transducers. The differences between this one-dimensional version and the two-dimensional extension of chapter 3 are described in Appendix C. The one-dimensional calculation will now be compared with experimental data for strips normal to a **Z** propagating wave and with other theoretical results.

The reflectivity of a single, isolated metal strip was computed and is given in fig. 4.16. The result is almost identical to results of a program by Milsom [29] (not shown), which uses accurate values for the effective permittivity function, taking bulk wave generation into account. It is concluded that bulk wave coupling has negligible effect at least in this case. Experimental results [40,32] are also given.

The reflectivity of strips with capacitative interactions was also computed for open and shorted strips and is shown in fig. 4.17, together with experimental results [24,25,10].

Further experimental results were obtained from the devices described in Section 1 of chapter 2, using shorted or open transducers as the strips, and these are also given. The experimental values are, on the average, smaller than the computed ones. The reason for this is not known.

A striking result ensues when reflectivity of strips with the following parameters is calculated. Fractional metallisation is kept constant at w/p=0.5, and period **p** is varied from 0 to 2**λ**. When **p** increases beyond **λ**, the potential of an open strip abruptly drops to zero, and remains at zero when **λ<p<2λ**. Since there is then no difference between open and shorted strips in this range, the reflectivity and velocity change due to shorted and open strips are identical (figs. 4.18 and 4.19). This phenomenon has not yet been explained, but the abrupt change in potential is apparent only when interactions between a large number of neighbouring strips is taken into account. The assumption that capacitative interactions are short ranged is obviously not valid in this case. The results of fig. 4.18 are in agreement with Datta and Hunsinger's [9], which use the same principles, but have an analytical method to calculate charge, rather than the numerical one used here.

4.9 Reflections through 90° by Strips Inclined at 45°

In the course of developing the dot reflection model, a version for reflections by strips through 90° was also produced. The strips are inclined at 45° to the incident beam if the material is isotropic. Details appear in Appendix D.

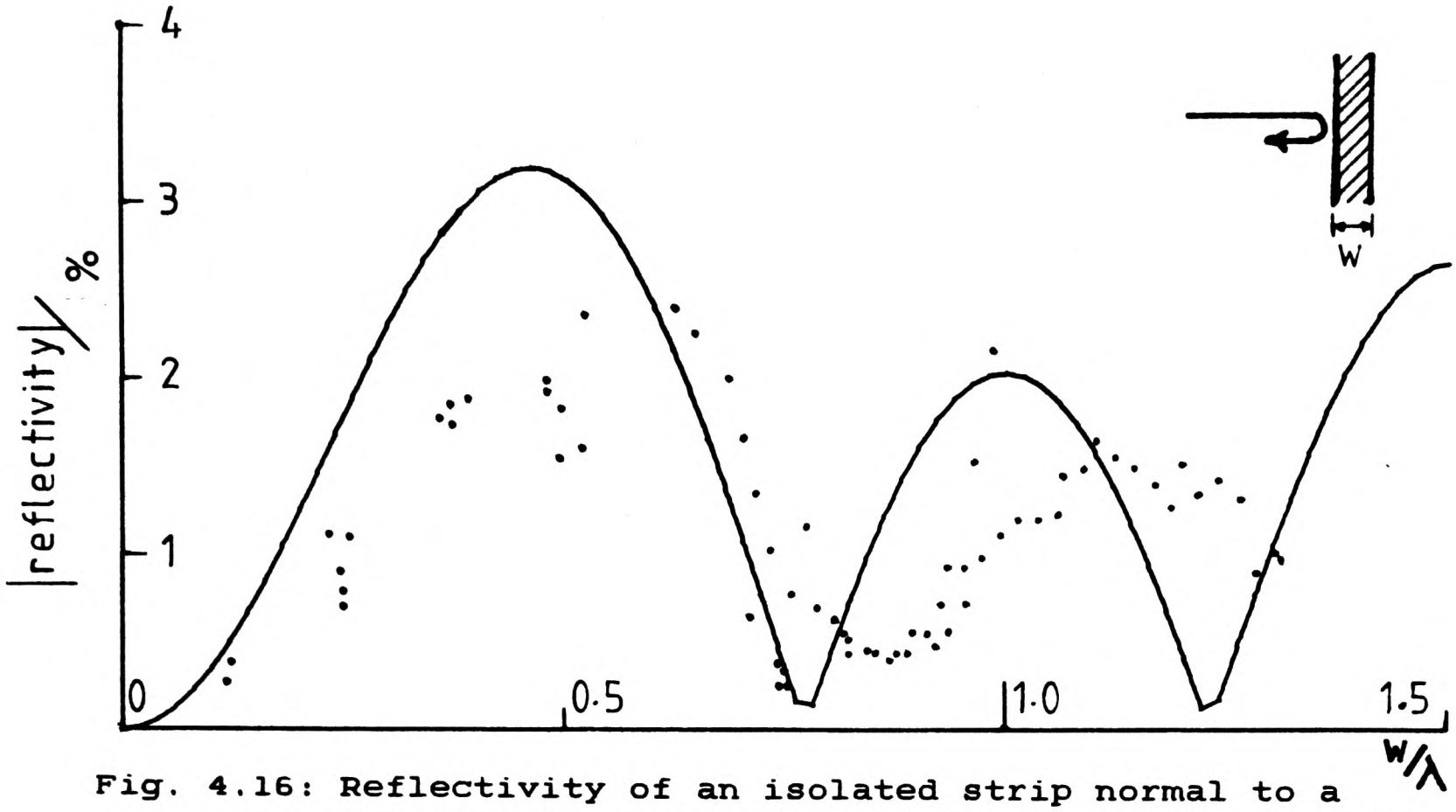


Fig. 4.16: Reflectivity of an isolated strip normal to a Z-propagating wave (magnitude only). Theoretical curve is almost identical to exact calculation [29]. Experiments from [40,32]

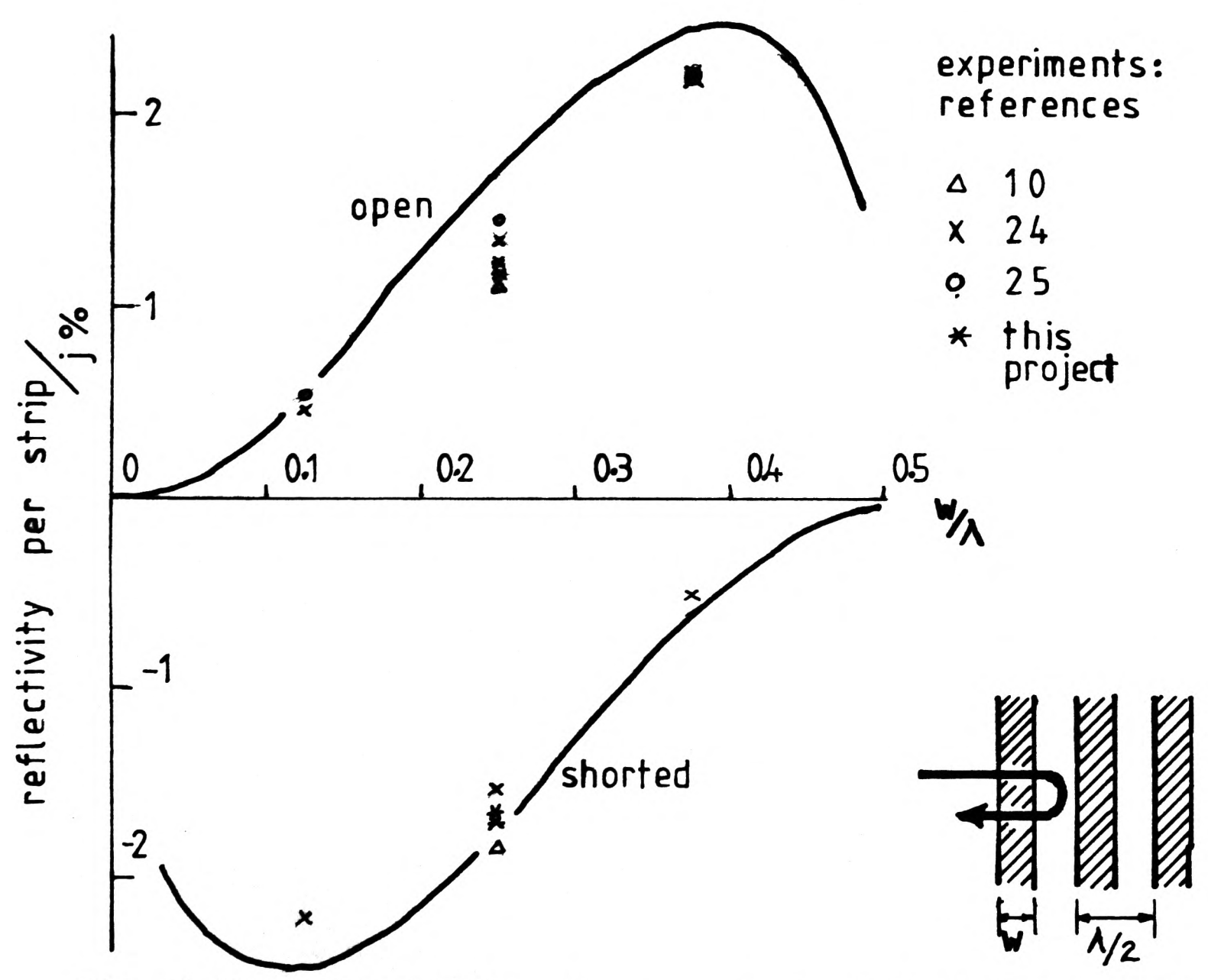


Fig. 4.17: Reflectivity of open and shorted strips normal to a Z-propagating wave (period constant; =0.5λ). (In some of the experiments, only magnitude was measured.)

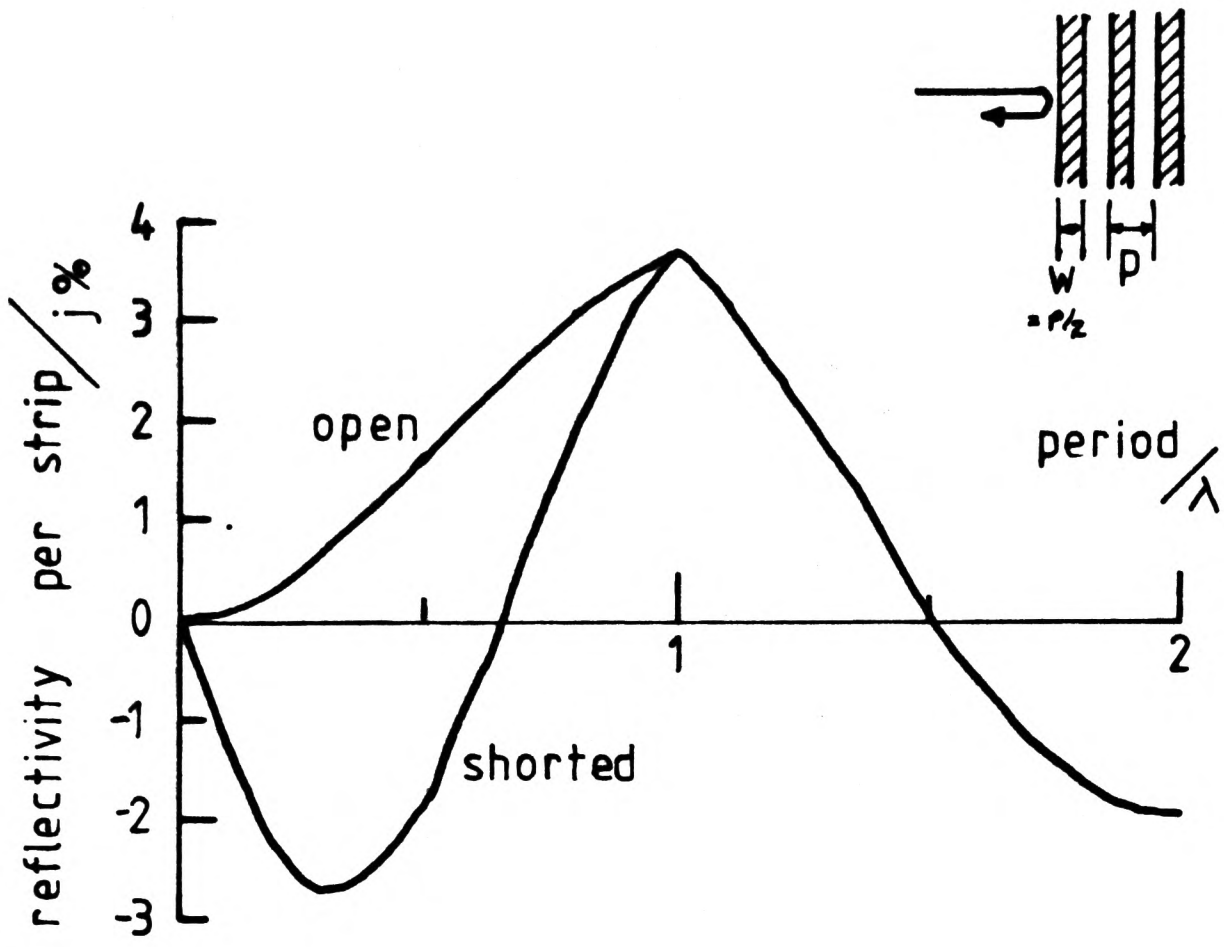


Fig. 4.18: Reflectivity of open and shorted strips normal to a Z-propagating wave (metallisation constant; =0.5).

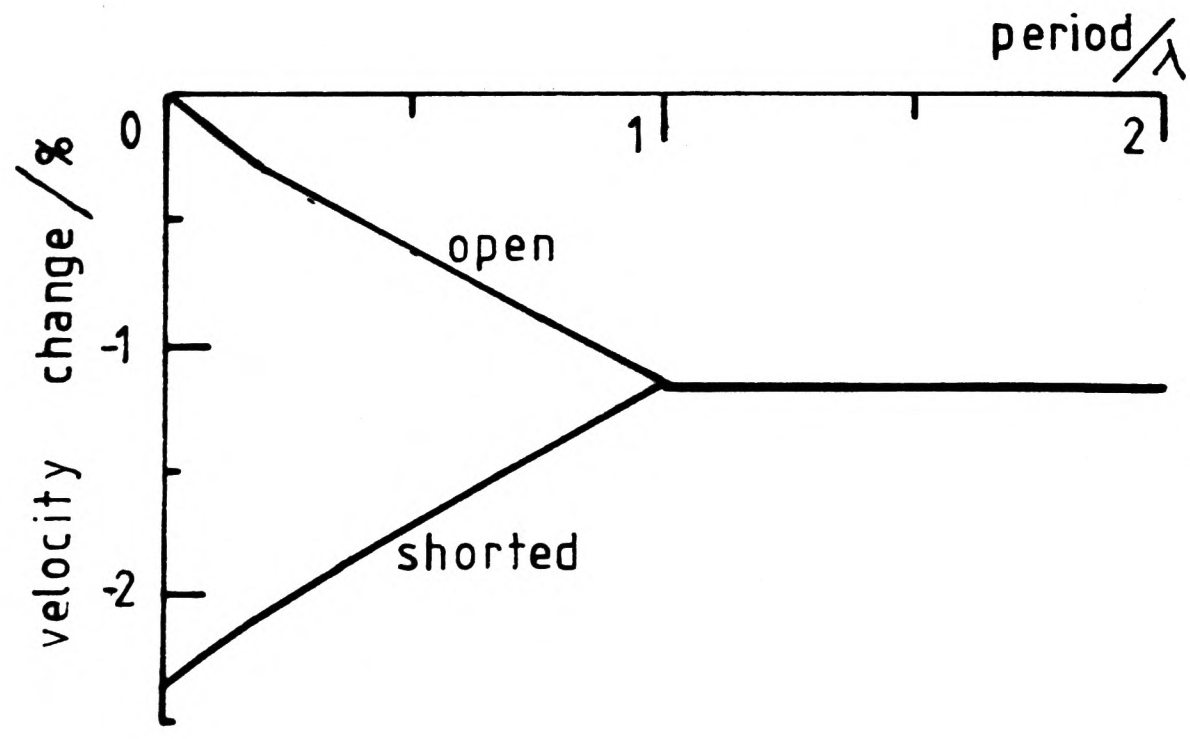


Fig. 4.19: Velocity change due to strips normal to Z-propagating wave (metallisation constant; =0.5).

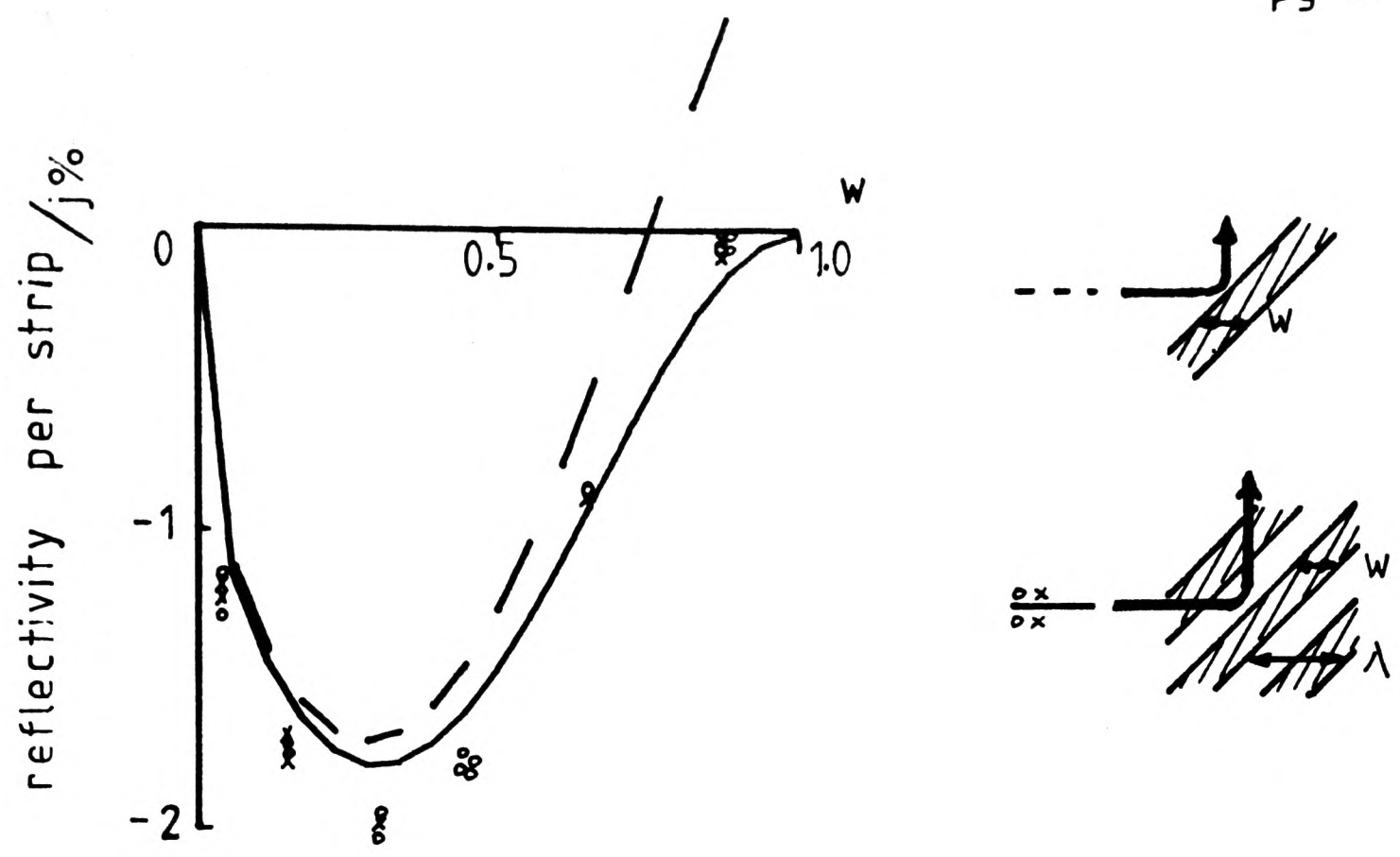


Fig. 4.20: Reflectivity of a strip at 45 degree incidence.
(-------): Isolated strip, theory x 0.89
(__o__x_): Strip with near-neighbours;
experimental data [19] and theory x 0.89

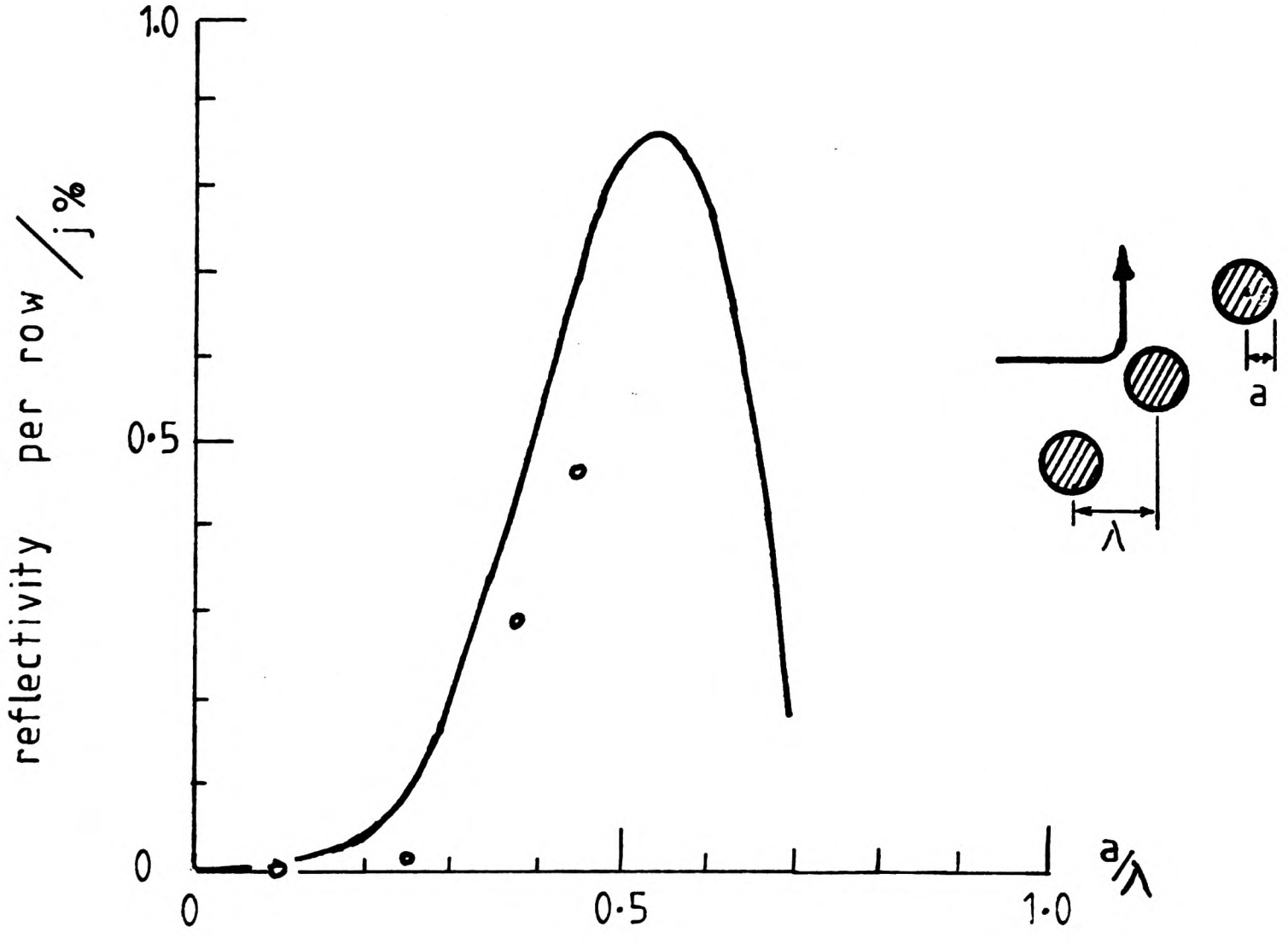


Fig. 4.21: Reflectivity by round dots.
(_____): Theory x 0.89 [33]; neglects near-neighbours
(o o): Experiment [17]; near neighbours present

Fig. 4.20 gives the results for isolated strips and reflection per strip for strips spaced one wavelength apart. The latter curve is compared with the results of Kitano and Nishikawa [19], showing good agreement. The theoretical value is slightly too small, probably because of the high metallisation (see section 5 and 6).

Bertoni [2] obtains analytically a curve for reflections by strips in a periodic array, which has the same shape as the result obtained here. However, the constant of proportionality for Y-cut $LiNbO_3$ reflecting from **Z** to **X** was not evaluated. Bertoni's result was not available when this work began.

4.10 Reflections from Circular Dots

The charge density on a circular dot can be solved analytically using spherical Bessel functions. This has been done by Peach [33]. The reflected wave is evaluated from charge using (3.20). Using the notation of chapter 3, a row of round dots has reflectivity

$$jr = -j\sqrt{G_z G}\,\varepsilon_\infty \frac{\alpha}{\pi} \left\{ 2\,\mathrm{sinc}^2\alpha - \mathrm{sinc}(2\alpha) - 1 - 2\,\alpha^2 \sum_{n=1}^{\infty} \cos(n\theta)\left[J_{n-1}^2(\alpha) - J_{n-2}(\alpha)\,J_n(\alpha)\right] \right\} \qquad (4.2)$$

where

$$\alpha = \frac{2\pi a}{\lambda} \qquad (4.2a)$$

θ is the angle between the reflected wave and **Z**

J_n are spherical Bessel functions

L (3.20) is normalised to λ

In fig. 4.21, this relation is compared with experimental data by Kitano and Nishikawa [17], extrapolated to zero metal thickness. Interactions are neglected in equation (4.2), but are present in the experimental data. By comparing square dots with and without nearest neighbours, it was estimated that the effect of the nearest neighbours is to reduce reflectivity by up to 20%. When this is taken into account, theory and experiment agree reasonably well.

It is interesting to note in passing that, from (4.2), the reflectivity of very small dots can be shown to increase with the fifth power of radius. Small, square dots (a=b) also have this property, not proved here. However, this information is unlikely to be useful since reflections by mass loading are likely to dominate for dots of this size.

CHAPTER 5

THE DESIGN OF A PULSE COMPRESSOR

The parameters required for the design of dot array devices are reflectivity, velocity change and scattering properties of the dots. These have been studied extensively, and the results were given in chapter 4. The final part of this project is to design a device using this information, since this is the best way to find out whether there are any further difficulties to be overcome.

The device chosen is the radar pulse compressor, or RAC (Reflective Array Commpressor) described in chapter 1. The specifications are typical of what might be required in a high-performance radar system: a bandwidth of 40 Mhz and dispersive delay of 25 microseconds, giving a time-bandwidth product of 1000. The centre frequency was chosen to be 160 MHz, so that the fractional bandwidth was 25%, a convenient value for the transducers. The device was designed to be down chirp, that is, with group delay increasing as frequency decreases. Before the design procedure is described, a brief description of chirped signals, the corresponding matched filters and their implementation in RACs will be given. The account deals mainly with linear FM chirps, in which the

instantaneous frequency varies linearly with time.

A general chirped signal can be expressed as

$$s_r(t) = a(t)\cos[\theta(t)] \tag{5.1}$$

or

$$s_r(t) = \tfrac{1}{2}\left[\, a(t)\exp[j\theta(t)] + a(t)\exp[-j\theta(t)] \,\right] \tag{5.2}$$

where **a**(t) is the envelope function. For simplicity, only the first term in the brackets is considered in the relations that follow. The analysis for the sum of the two terms is similar. Thus a "partial" signal defined as

$$s(t) = a(t)\exp[j\,\theta(t)] \tag{5.3}$$

is used in the analysis.

5.1 The Principle of Stationary Phase

The principle of stationary phase is very useful for signals with large time-bandwidth product, where it leads to accurate approximations for frequency response. It is described in detail by Cook and Bernfeld [8], and only an outline is given here.

The Fourier transform of the signal

$$s(t) = a(t)\exp[j\,\theta(t)] \tag{5.4}$$

which has intantaneous frequency $\omega_i = \theta'(t)$ is

$$S(\omega) = \int_{-\infty}^{\infty} a(t)\exp[j(\theta(t) - \omega t)]\,dt \tag{5.5}$$

The real parts of **s**(t) and the integrand of (5.5) are sketched

in fig. 5.1 . At any frequency ω, the exponential factor in the integrand is stationary (as t varies) at time $t=\tau$ such that

$$\frac{d}{dt}\left(\theta(t) - \omega t\right)_{t=\tau} = 0$$

$$\text{or} \quad \theta'(\tau) = \omega$$
$$\text{i.e.} \quad \omega_i = \omega \qquad (5.6)$$

At other times, the intergrand is oscillating rapidly as shown in fig. 5.1(ii), and does not contribute significantly to the integral since adjacent positive and negative sections cancel. An approximate value of $S(\omega)$ can therefore be found by integrating over a small range of t near $t=\tau$, instead of over an infinite range or over the duration of the signal. By assuming that a(t) is constant over this range, and expanding $\theta(t)$ as a Taylor series about $t=\tau$, it can be shown that

$$S(\omega) = \sqrt{2\pi}\,\frac{a(\tau)}{\sqrt{\theta''(\tau)}}\exp\left[j\left(-\omega\tau + \theta(\tau) + \frac{\pi}{4}\right)\right] \qquad (5.7)$$

where τ is given implicitly by

$$\theta'(\tau) = \omega \qquad (5.7a)$$

So $S(\omega)$ depends only on the small region (or regions) of s(t) which has instantaneous frequency $\omega_i=\omega$. As will be seen later, this simplifies the design equations.

5.2 Chirped Signals and their Matched Filters

The pulse compressor is an example of a matched filter. A signal (such as a radar echo), which is affected by additive white gaussian noise, is detected with maximum signal-to-noise ratio if the receiver contains a matched

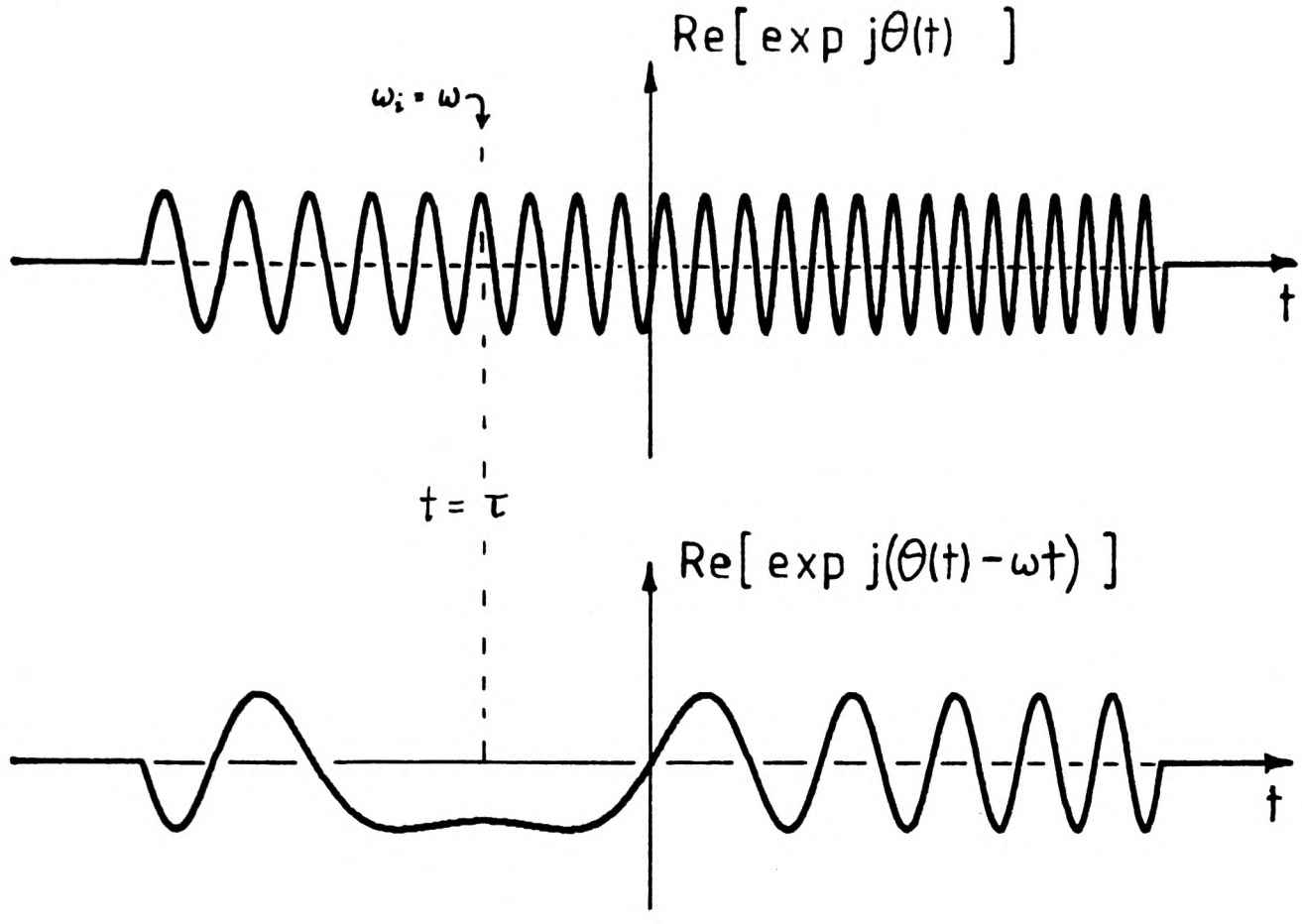


Fig. 5.1: Evaluation of a Fourier Transform using the Principle of Stationary Phase.

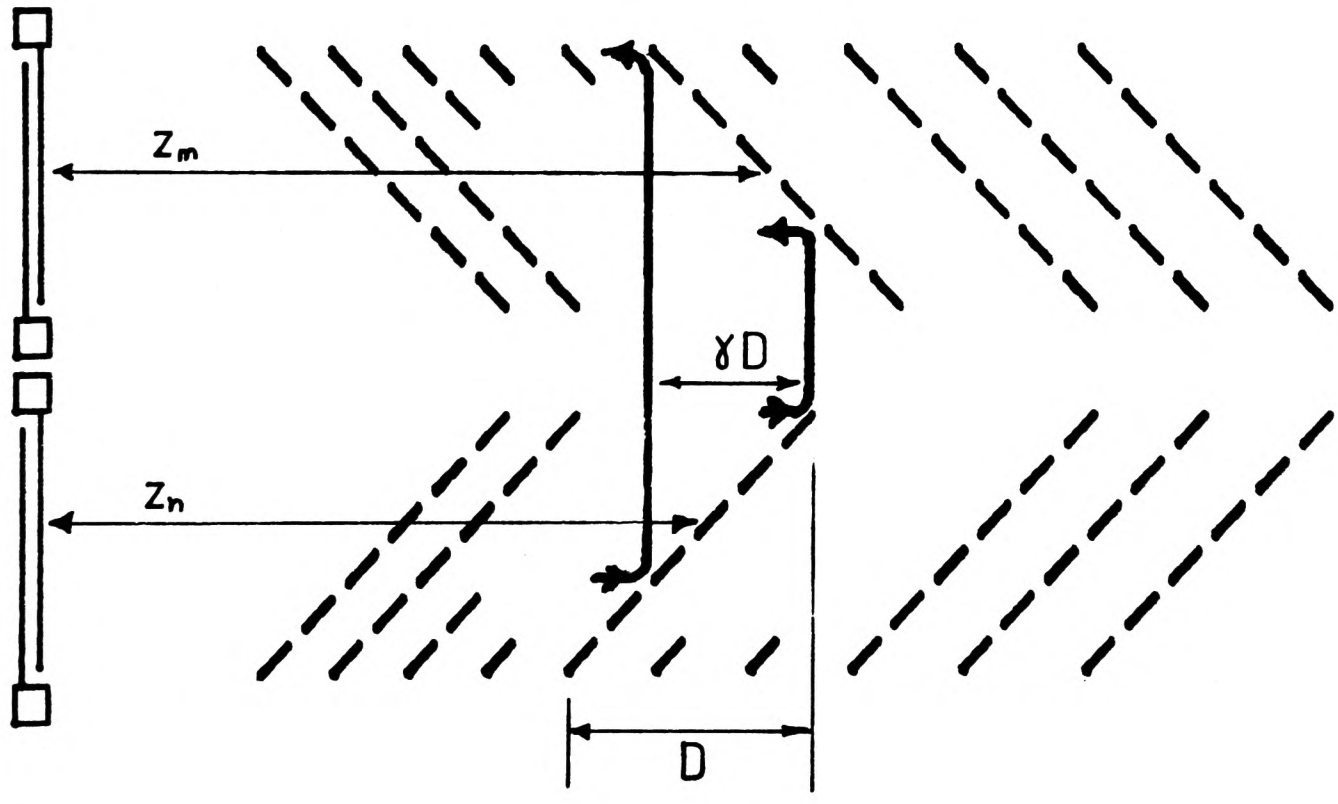


Fig. 5.2: Notation used in calculation of the frequency response of a RAC.

filter. If the signal is **s**(t) and has a frequency spectrum **S**(ω), the matched filter impulse response **f**(t) and frequency response **F**(ω) are given by

$$f(t) = K\, s^*(t_o - t) \tag{5.8}$$

$$F(\omega) = K \exp(j\omega t_o)\, S^*(\omega) \tag{5.9}$$

where the delay term t_o can be chosen arbitarily and will be taken to be zero for simplicity. **K** is an arbitary real constant. The asterisk denotes a complex conjugate. It is superfluous in $s^*(\,)$ for real signals, but is necessary for the "partial" complex signals used here.

A linear FM up-chirp signal has the form

$$s(t) = a(t) \exp\left[j\left(\omega_o t + \tfrac{1}{2}\mu t^2\right)\right] \tag{5.10}$$

where μ is positive and **a**(t) is a rectangle function

$$\left.\begin{aligned} a(t) &= 1 \qquad -\tfrac{T}{2} < t < \tfrac{T}{2} \\ a(t) &= 0 \qquad \text{elsewhere} \end{aligned}\right\} \tag{5.10a}$$

The rectangular envelope is the function usually imposed by the high-power transmitter. The instantaneous angular frequency of **s**(t) is

$$\omega_i = \omega_o + \mu t \tag{5.11}$$

The Fourier transform of **s**(t) is found from (5.7) and (5.7a).

$$S(\omega) = \sqrt{\frac{2\pi}{\mu}}\, a\left(\frac{\omega-\omega_o}{\mu}\right) \exp\left[j\left(-\frac{(\omega-\omega_o)^2}{2\mu}+\frac{\pi}{4}\right)\right] \qquad (5.12)$$

Subject to the limitations of the stationary phase approximation, $S(\omega)$ has constant magnitude and quadratically varying phase inside the frequency range

$$|\omega-\omega_o| \leq \mu T/2 \qquad (5.13)$$

and is zero outside this range. If $\Delta\omega_i$ is defined as

$$\Delta\omega_i = \omega_{i(max)} - \omega_{i(min)} \qquad (5.14)$$

then the frequency range is

$$|\omega-\omega_o| \leq \frac{\Delta\omega_i}{2} \qquad (5.15)$$

The impulse response $f(t)$ of the ideal matched filter for the signal $s(t)$ is given by (5.8) together with (5.10) and its frequency response by (5.9) and (5.12).

$$f(t) = \sqrt{\frac{\mu}{2\pi}}\, a(-t) \exp\left[j\left(\omega t-\tfrac{1}{2}\mu t^2\right)\right] \qquad (5.16)$$

$$F(\omega) = a\left(\frac{\omega_o-\omega}{\mu}\right) \exp\left[j\left(\frac{(\omega-\omega_o)^2}{2\mu}-\frac{\pi}{4}\right)\right] \qquad (5.17)$$

The value of K is taken to be $\sqrt{(\mu/2\pi)}$ so that $|F(\omega)|=1$ in the pass band, and noise power is unaffected by the filter. The frequency spectrum of the signal after matched filtering is

$$S_o(\omega) = S(\omega)\, F(\omega)$$

$$= \sqrt{\frac{2\pi}{\mu}}\, a\left(\frac{\omega-\omega_o}{\mu}\right) \qquad (5.18)$$

which is a constant-phase rectangle function. Taking inverse

Fourier transforms,

$$s_o(t) = \sqrt{BT}\,\mathrm{sinc}(\pi Bt)\,\exp(j2\pi f_o t) \qquad (5.19)$$

where

$$B = \frac{\Delta\omega_i}{2\pi} \qquad (5.19a)$$

The envelope of $s_o(t)$ is

$$a_o(t) = \sqrt{BT}\,\mathrm{sinc}(\pi Bt) \qquad (5.20)$$

which has a -4dB pulse width of 1/**B**. So pulse width is compressed by a factor of **TB** from **T** to 1/**B**. Peak amplitude increases by a factor of √(**TB**). Since the frequency respose of the filter has been normalised to pass the noise power unaffected, the peak-signal to noise ratio (of powers) has been increased by **TB**.

The **sinc** function implies time-sidelobes (subsidiary peaks) of -13.3 dB. In a radar system, sidelobes are a disadvantage because they can mask a nearby smaller echo. They can be reduced by modifying the design of the filter, at the expense of reducing the S/N ratio and increasing the pulse width, that is, by deviating slightly from the matched filter condition. One method of reducing sidelobes is to weight the frequency response of the filter. The "Hamming" weighting, for example,

$$\left|F(\omega)\right| = 0.54 + 0.46\cos\left(2\pi\frac{\omega-\omega_o}{\Delta\omega_i}\right) \qquad (5.21)$$

cuts sidelobes to -42.8 dB while widening the pulse width by a factor of 1.47 and introducing 1.34 dB of mismatch loss [8]. This can be confirmed by modifying (5.17) to have this

amplitude weighting, and then evaluating $s_o(t)$. Since a small region of the impulse reponse is related to each frequency in the frequency domain by the principle of stationary phase, the weighting can equivalently be done in the time domain. (If the **TB** product is small, so that the principle of stationary phase is not a good approximation, it has been shown [26] that weighting in the time domain is preferable as far as sidelobe reduction is concerned.)

5.3 Standard RAC Equations

The RAC is one implementation of the pulse compressor for chirped signals. Equations relating specifically to RACs are now given. In the notation used here, the length of the RAC is parallel to the **Z** axis. The following equations (5.22,24,25) are obtained from Williamson [47] with slight changes in notation. The position of the **n**th strip is given by

$$Z_n = Z_o + \frac{V f_1 T}{2B}\left[1 - \sqrt{1 - \frac{4B}{f_1^2 T} n}\right] \qquad n = 0, 1, \ldots \frac{T f_o}{2} \qquad (5.22)$$

where

V = Velocity of Z-propagating wave, assumed constant

f_1 = Highest instantaneous frequency

T = Duration of the signal (5.22a)

B = Bandwidth of the signal.

The spacing between strips is given by

$$\Lambda_{i-\frac{1}{2}} = Z_i - Z_{i-1} \simeq \left(\frac{dZ}{dn}\right)_{i-\frac{1}{2}} = \frac{V}{f_1}\left[1 - \frac{4B}{f_1^2 T}\left(i - \tfrac{1}{2}\right)\right]^{-\frac{1}{2}} \qquad (5.23)$$

Calculation of frequency response is illustrated in fig. 5.2 . The contribution from a pair of reflectors, one in each array, is given by their reflectivities, the total path length, and the overlap factor γ between the two reflectors. The frequency response of the arrays is found by summing over all reflectors in the arrays:

$$F(\omega) = \sum_{m,n} r_m r_n \gamma_{mn} \exp\left[-j\omega \frac{Z_m + Z_n}{V}\right] \qquad (5.24)$$

where the length of the path in the **x** direction is neglected since it only leads to a linear phase term.

An approximate version of (5.24) was developed by Martin [23] by index transformation and using the principle of stationary phase. When the positions z_m, z_n are defined as in (5.22), (5.24) reduces to

$$F(\omega) = \left(N_e(\omega)\, r(\omega)\right)^2 \; g\left(\frac{D}{N_e \lambda}\right) \; \exp\left[-j\left(\omega\tau_o - \frac{T(\omega-\omega_b)^2}{2\Delta\omega_i}\right)\right] \qquad (5.25)$$

where

$$N_e = \frac{\omega}{\Delta\omega_i}\sqrt{\frac{T\Delta\omega_i}{4\pi}} = \frac{f}{B}\sqrt{\frac{TB}{2}} \qquad (5.25a)$$

τ_o is the group delay at centre frequency. $g(D/N_e\lambda)$ is a complex function. However, in the device to be designed, **D** (equal to aperture for an isotropic substrate) is constant and $N_e\lambda$ can be seen to be constant from (5.25a), so **g**( ) is constant. $jr(\omega)$ is the reflectivity* of those reflectors which are synchronous at frequency ω. The frequency response has quadratic phase, which by comparison with (5.17) verifies that (5.22-25) describe a matched filter for a linear FM signal if

* per reflector

the amplitude function $|N_e r(\omega)|^2$ is chosen appropriately. However, for reasons already discussed, the amplitude function of the device that will be designed will be Hamming weighted.

The attenuation of the SAW on $LiNbO_3$ has been found experimentally by Szabo and Slobodnik [42]. Allowance for this attenuation will be made using a simplified version of the relation given, valid near 160 MHz:

$$0.19\left(\frac{f}{f_o}\right) + 1.06\left(\frac{f}{f_o}\right)^2 \quad dB/\mu s \tag{5.26}$$

where

$$f_o = 1000 \text{ MHz} \tag{5.26a}$$

One final standard publication on RACs will be mentioned. Wright and Haus [49] treat multiple reflections in RACs analytically using confluent hypergeometric functions. It was earlier thought that the relations were potentially very useful. It was eventually learned that, in the dot array design chosen, unwanted reflections at a given frequency occured along several Bragg lines in different parts of the array, so that these analytic relations would be difficult to apply directly. Instead, a simple, approximate method of allowing for these double reflections was developed, and will be discussed later.

5.4 Considerations in the Design of a Dot RAC

The design procedure for groove RACs is well established, but when designing a dot RAC, several new considerations arise. For example, the dots form several Bragg lines, that is, an array can have several scattering vectors. Since the

reflection mechanism in the proposed device is electrical, coupling to the leaky wave is a second possible problem, as the electrical coupling coefficient of the LSW is high. In this respect, the proposed device differs even from previous dot array devices [34,36-39,43,44] which employ mass loading. The benefit of experience is not available for these new considerations, but as far as possible, the difficulties are identified and allowed for.

These differences are more significant than they might first appear. As explained in chapter 1, fabrication of masks is often easier if there are a relatively small number of dots. Dot size is therefore chosen such that the reflectivity per dot is relatively large, so that there need be only a small number of dots per row. The number of possible scattering vectors is therefore large. The power loss due to scattering of the SAWs can obviously be significant. The large number of scattering vectors also increases the possiblility of mode conversion to bulk waves and LSWs, a problem made worse by the high LSW coupling. In addition, the scattered waves can reach the output transducer, possibly forming a spurious reflection path.

Previous dot devices had randomised dot positions, in which the positions of the rows were defined by the required response, but the position of each dot along the row was random. As related in Appendix I, the effect is to spread the power loss over a wide band of frequencies, thereby avoiding notches in the reponse. When the power loss is averaged over frequency, the effect of randomisation is slight. In the

proposed device, it seems necessary to avoid spurious reflections by having a good array pattern rather than randomising the positions of the dots.

As the reflectivity of dots as a function of size is now available, there are two alternative methods of weighting. One method is by the number of dots per row. In the device to be designed, there are only about 10 dots per row, so that the quantisation error after two reflections would be ±10%. There is a disadvantage in simulating (for example) 10.2 dots per row by having sets of five rows with 10,10,11,10,10 dots respectively: this would complicate the array pattern and create additional scattering vectors. It was therefore decided that weighting would be by dot size. Dot "length" **a** would be varied according to the weighting required. To reduce scattering, **a** would have to be greater than about $0.7\lambda_i$ over most of the array (see chapter 4, section 4). Dot "width" **b** would be kept at $0.2\lambda_i$ so that reflectivity would be relatively insensitive to overetching, (chapter 4, section 3). (λ_i is the wavelength of the SAW synchronously reflected by the particular part of the array.) The distance between dots in a row, L_1, would be constant (not proportional to λ_i) because this greatly simplifies the generating GAELIC program. The simplification results from the fact that each array can be made of several identical sections (fig. 5.3).

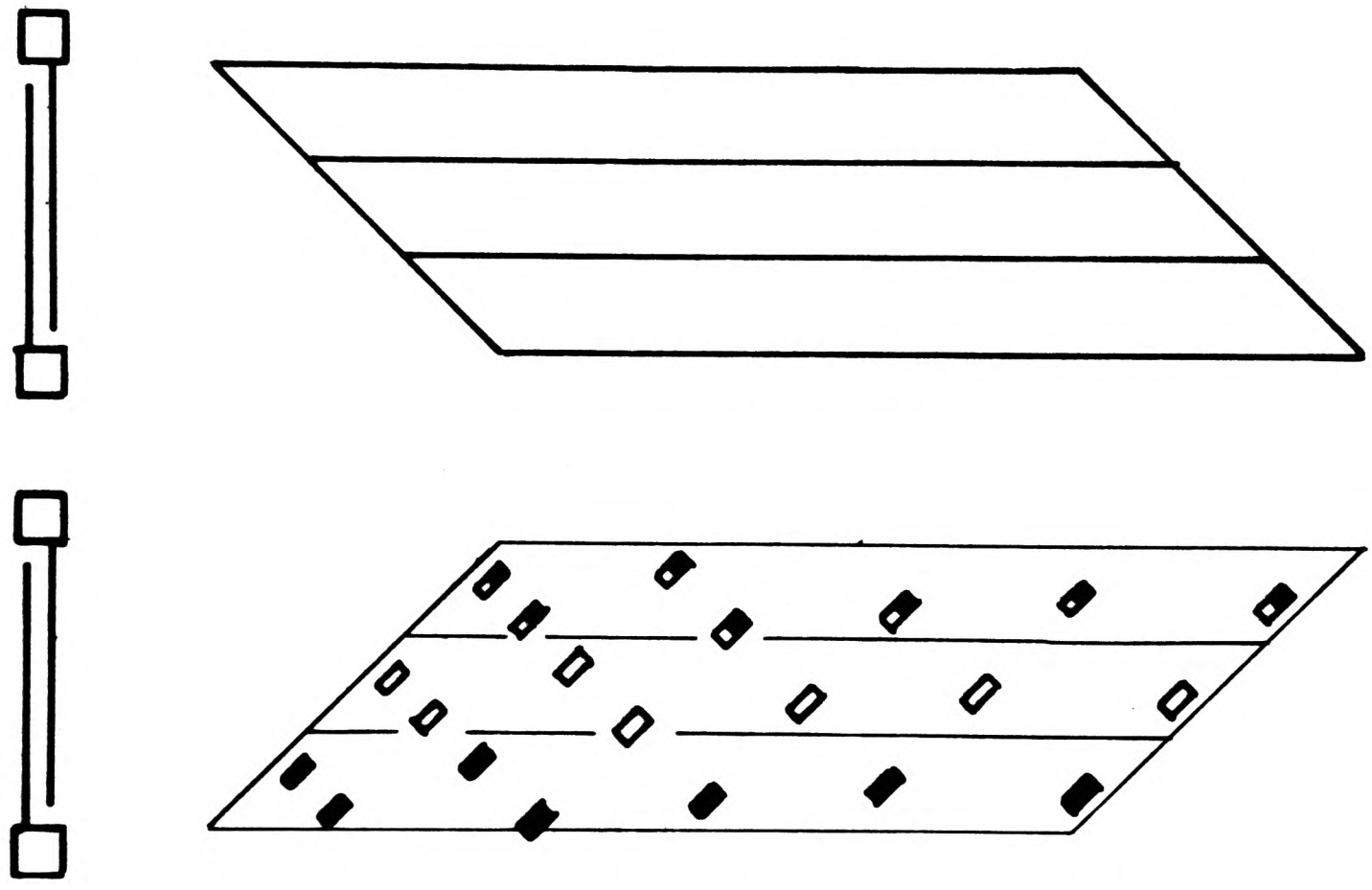


Fig. 5.3: Identical sections forming an array, showing that a step-and-repeat process is possible.

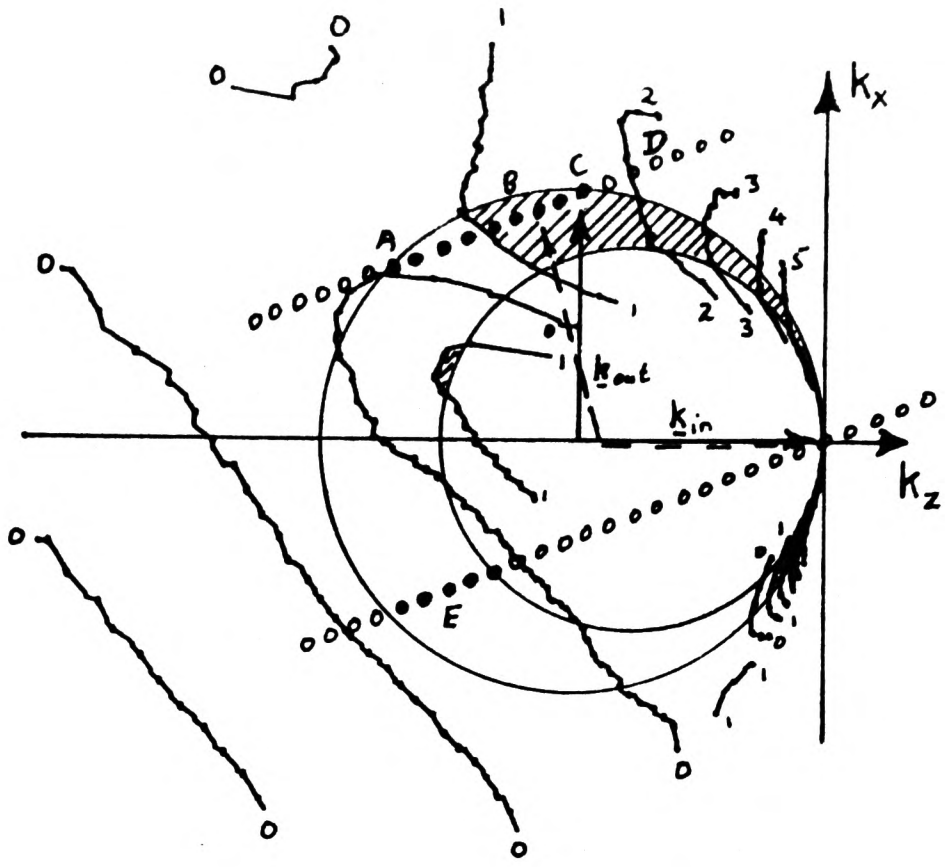
5.5 Features of the Selected Array Pattern

The array pattern selected is illustrated in fig. 5.4 . $L_o=8\lambda_o$ was selected to give a suitable level of reflectivity. (λ_o=wavelength at centre frequency.) The angle of inclination of the rows with the **Z** axis is chosen so as to reflect the beam through 90° on the anisotropic substrate; the calculations involved are described later. The inclination of the "columns" is not critical. It was originally taken to be $\tan^{-1}$ 2.5, but after scaling to allow for velocity anisotropy ($v_z \neq v_x$), it becomes 69.4°. It is chosen to reduce the spurious reflections, as described below.

5.5.1 Scattering of SAWs by the Array

The scattering vector corresponding to reflection through 90° is independent of the inclination of the "columns", which can therefore be selected so as to reduce scattering by other scattering vectors. The basis of this scheme has been described by Bloch, Paige and Stove [3]. An additional consideration is that the dots are not isotropic scatterers, so that scattering vectors in certain directions can be neglected. It is re-emphasised that the term "scattering", when applied to arrays of dots, refers to **coherent** reflections by the array in undesirable directions.

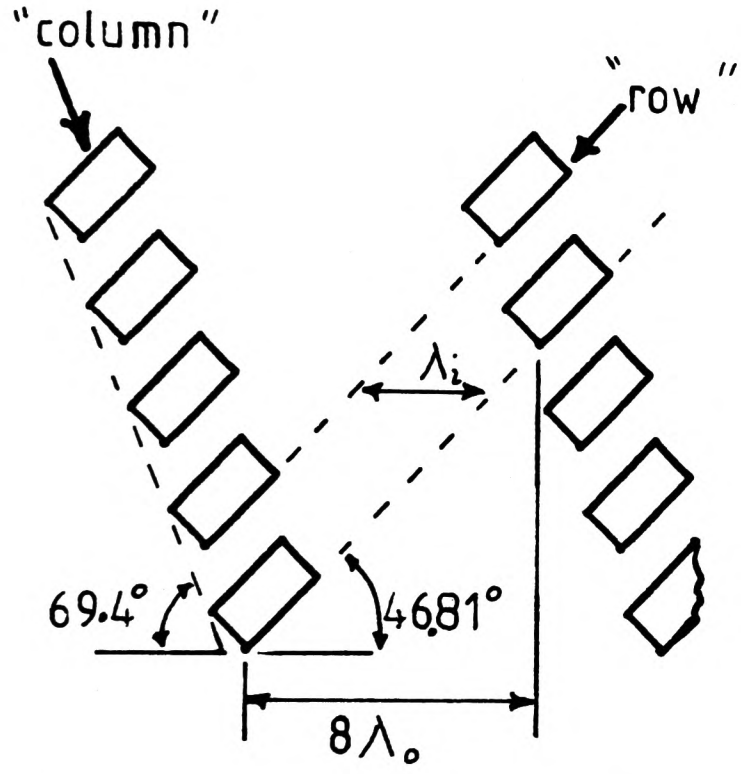
A short section of the array (shaded in fig. 5.5), in which period is approximately constant at λ_i, can be assumed to be periodic. The Fourier transform $H(k_z,k_x)$ of the array pattern (or "array factor") determines the scattering vectors as described in chapter 3. $H(-k_z,-k_x)$ is given in



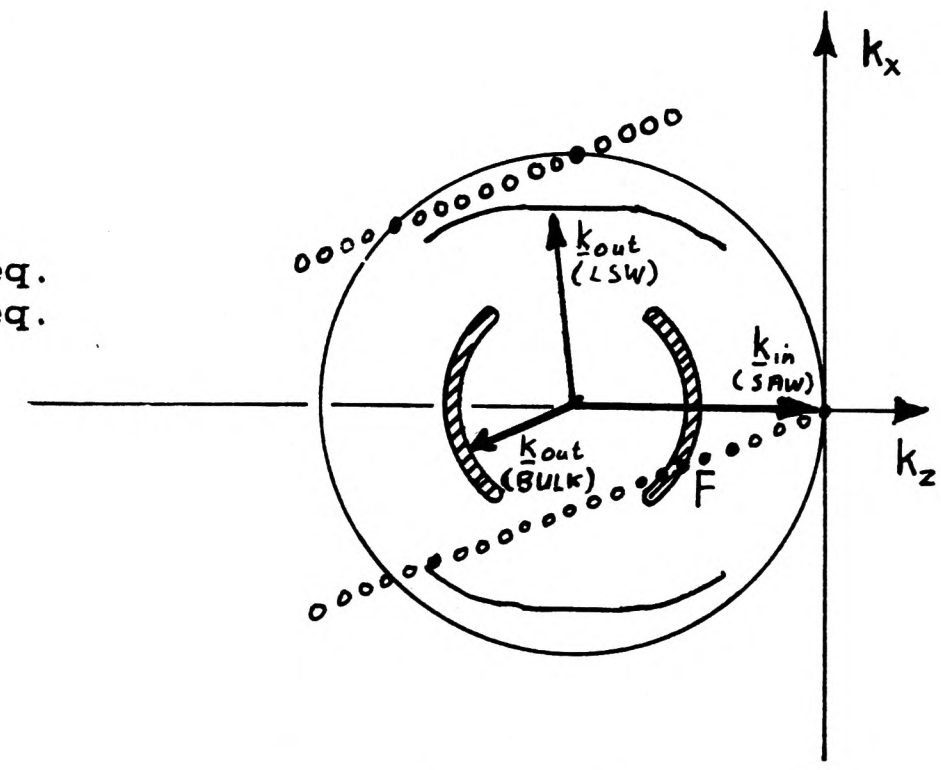
(ii) 1. Fourier transform of a segment of the array pattern (points, representing delta functions), superimposed on

2. Contours of dot scattering coefficient ($a=\lambda_i$, $b=0.2\lambda_i$, near-neighbours as shown in (i), and

3. Ewald circles



(i) Dot pattern selected for RAC.
λ_i = wavelength at synch. freq.
λ_o = wavelength at centre freq.



(iii) Ewald 'circles' for SAW, LSW and Bulk Waves.

Fig. 5.4:

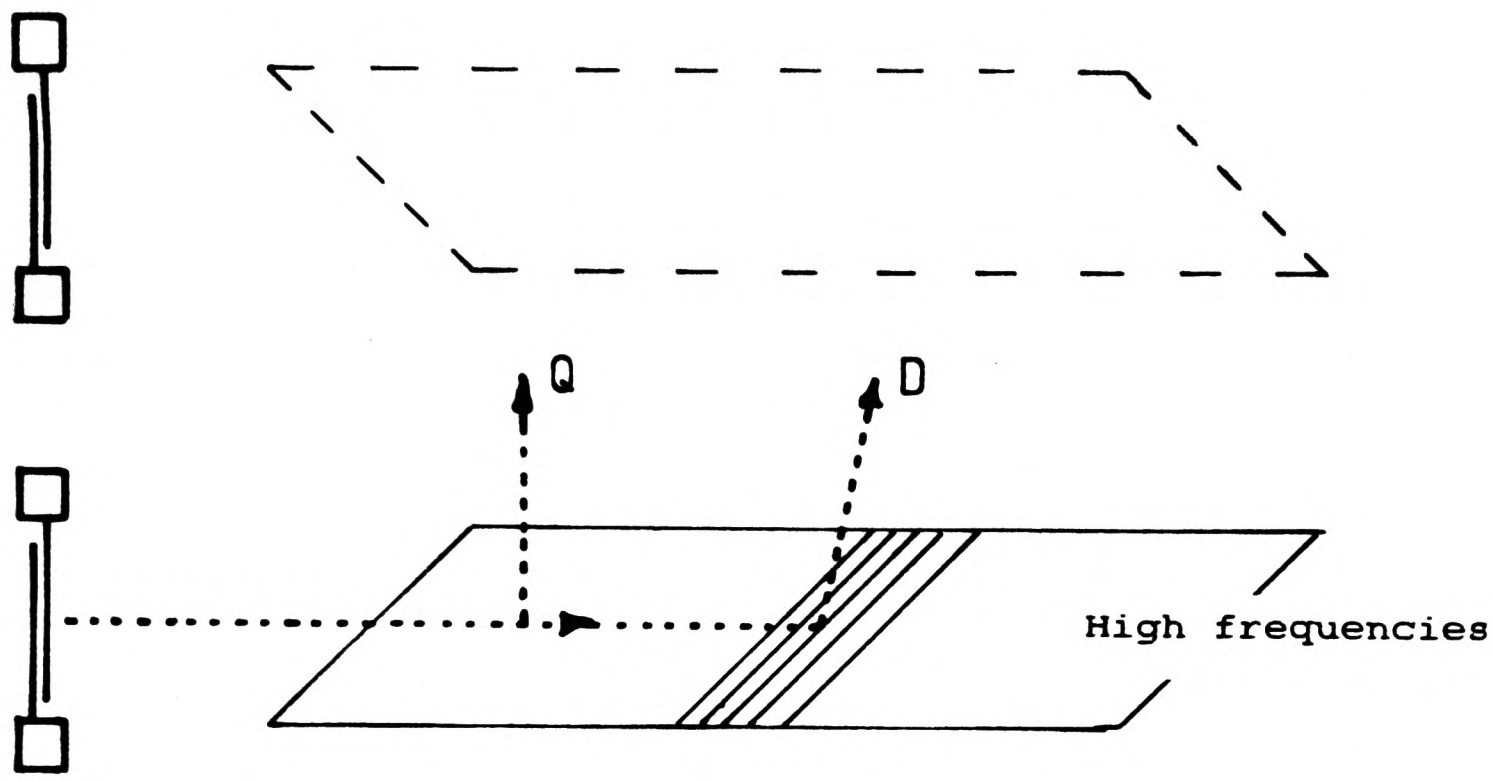


Fig. 5.5: Reflections by various scattering vectors at various frequencies.

fig. 5.4(ii). The strength of the reflection corresponding to any scattering vector depends on the reflectivity of the dot in that particular direction and frequency. This can be read off the superimposed reflectivity contours for any particular dot size.

A section of the array which reflects frequency f synchronously through 90° must pass frequencies $f_1 < f$ with a minimum amount of attenuation, since waves of frequency f_1 are reflected further along the array in a down-chirp device. In fig. 5.4(ii), scattering vectors corresponding to the lower frequencies or interest are inside the crescent drawn. It can be seen that reflectivities are relatively small, particularly for scattering vectors near "E". Energy loss by the scattering vectors along ABC, being more significant, will be accounted for in the design. Reflection paths such as D (fig. 5.5) corresponding to scattering vectors D (fig. 5.4) cause a larger power loss, but these are irrelevant because the useful signal has already been diverted along Q.

A second condition is that scattered SAWs along $\overset{B}{\wedge}$ and D are not reflected by the second array towards the output transducer. This condition is seen to be satisfied, since the low-frequency signal B is reflected towards the high-frequency end of the second array, and the high-frequency signal D is directed towards the low-frequency end.

Finally, there is the possibility that reflections such as A may reach the output transducer directly. In this array, the scattering coefficient is weak, and in addition, the reflection feeds a far-out spatial sidelobe of the transducer,

so again, the spurious signal should be small.

The same arguements apply for other sizes of dots ($b=0.2\lambda_i$, $a>0.7\lambda_i$). Although there is a large number of scattering vectors, their effect has been minimised by careful choice of array pattern.

5.5.2 Scatttering of Bulk and LSWs by the Array

Since a complete model for mode conversion by dots of a SAW to Bulk and LSWs has not yet been produced, their effects have to be reduced by the choice of array pattern rather than dot shape.

LSWs will be considered first. In order to satisfy the scattering condition

$$\underline{k}_{out} = \underline{k}_{in} + \underline{k}_g \qquad (3.25)$$

the tip of a scattering vector $\underline{k}_g$ must lie on the LSW slowness surface shown in fig. 5.4(iii). The slowness surface is obtained from velocity data given by Marshall and Paige [21]. Only part of the slowness surface is shown because coupling to LSWs (given by $(\frac{\Delta v}{v})_{LSW}$) is very weak near the Z direction. It can be seen that none of the scattering vectors touch the slowness surface, so that LSWs are not reflected coherently. The same applies to lower frequencies f_1.

The same arguments can be applied to Bulk wave generation, with $\underline{k}_{out}$ representing the component of the wave vector parallel to the surface [21]. The radiation resistance plots (fig. 4.15) indicate that for a given value of $\underline{k}$ defined by the transducer finger spacing, large Bulk wave

losses occur when the frequency ω is twice that required to generate SAWs. This implies that for a given frequency ω, the value of $\underline{k}_{out}$ corresponding to large reflections is about half the $\underline{k}_{out}$ for SAWs. The locus of $\underline{k}_{out}$ values for strong bulk wave coupling in various directions is also produced in fig. 5.4(iii). As with LSWs, only the part of the curve corresponding to strong coupling is drawn. The scattering vectors are clear of these curves except near F, so some reflections due to these grating vectors is expected. However, the power lost also depends on reflectivity per dot, and this is believed to be small in that particular direction. This inference was reached by assuming that the amplitude of the generated bulk wave is proportional to $\bar{\sigma}(\underline{k}_{out,bulk})$, which was calculated by one of the programs described in chapter 3 and found to be small.

5.6 Data on Dots and Transducers used in the RAC Design

The dots used in the design of the RAC differed slightly from any of the dots in the previous experiments in that the positions of the nearest neighbours were different, for reasons already discussed. However, a computer program was available to calculate the new values of reflectivity, velocity and scattering properties.

Scattering contours such as the set presented in fig. 5.4(ii) were obtained for dots with $b=0.2\lambda_i$, $a=0.2\lambda_i$ to $1.8\lambda_i$, and nearest neighbours as in the pattern indicated by fig. 5.4(i). 90° reflectivity jr was calcualted from this, and is given in fig. 5.6 as a function of a/λ_i. The graph is normalised to $L_1=\lambda_i$. Velocity perturbation is given as a

function of a/λ_1 and f/f_i (or λ_i/λ) in fig. 5.7, again normalised for $L_1=\lambda_i$. From the scattering diagrams (e.g. fig. 5.4(ii)) it was confirmed that the reflectivity associated with the scattering vectors marked E were small over most of the range of (a/λ_i) eventually used in the design. Since the reflectivities of the scattering vectors between A and C were larger, the average value of r_s^2 (where jr_s is the reflectivity of a row of dots in the direction of one of the scattering vectors along AC) was evaluated graphically. $(r_{os}^2)_{av}$ is plotted in fig. 5.6, also normalised. It is shown in Appendix G that power lost by the 90° reflection is approximately $(rN_e)^2$, even for narrow arrays. The loss associated with each of the other relevant scattering vectors is assumed to be also $(r_sN_e)^2$, with an appropriate value of r_s, since the angle through which the SAW is reflected is in the neighbourhood of 90°.

It is worth making a few comments on this data. In fig. 5.6, r_o is rapidly-varying for small values of a/λ_i, but this does not lead to large errors in weighting because weighting is proportional to r^2, being due to two arrays. A plot of r^2 (not shown) does not have this rapidly-varying section.

The fractional velocity perturbation is relatively large when compared with grooves, and is up to about 3×10^{-3} when $L_1=8\lambda_i$. This compares with less than 10^{-4} in a groove device [46] with similar specifications. It also varies with a over a wide range. The variation of velocity with frequency is much less significant.

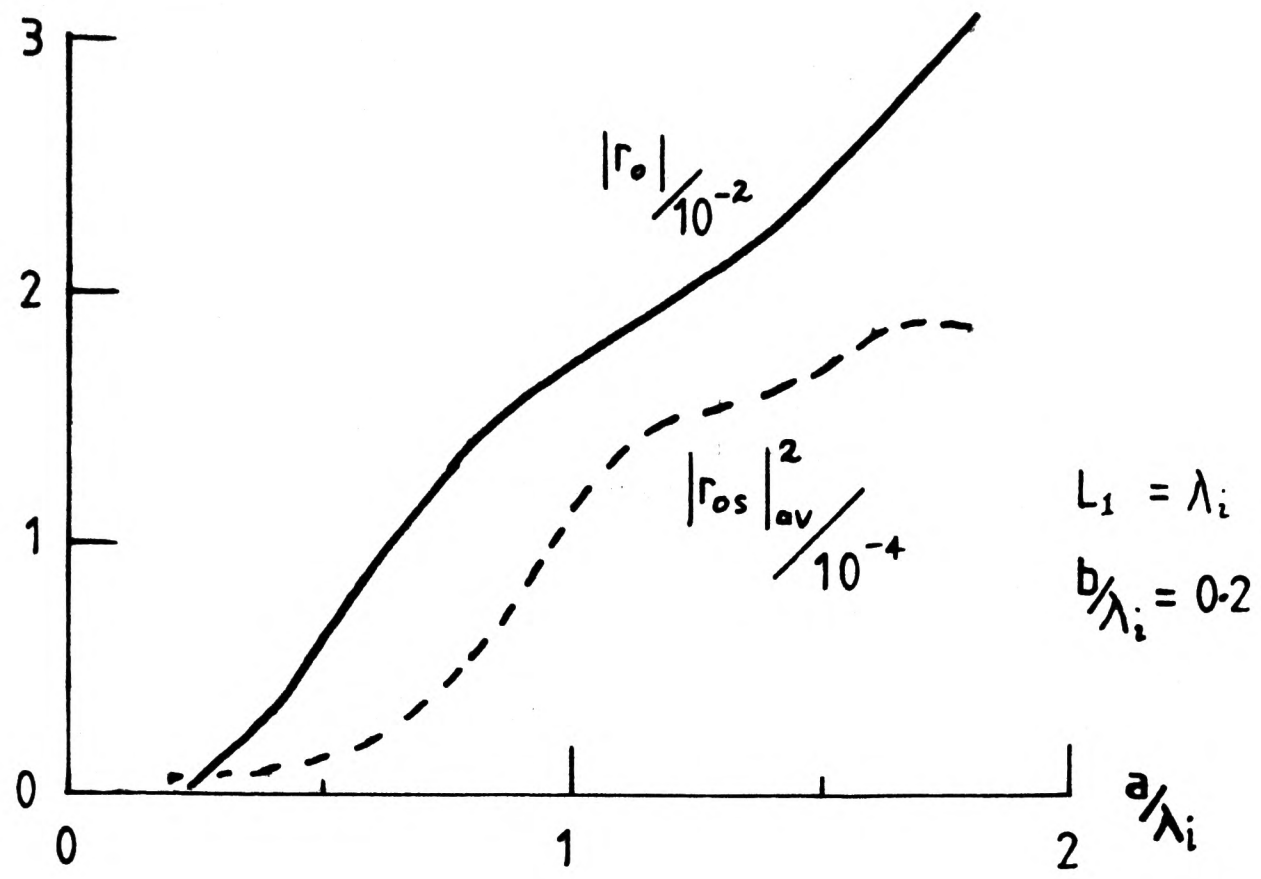


Fig. 5.6: (_____) Reflectivity per row of dots in the pattern of fig 5.4, normalised to $L_1 = \lambda_i$.
(----) Average value of r^2 for the scattering vectors A to C in fig. 5.4(ii).

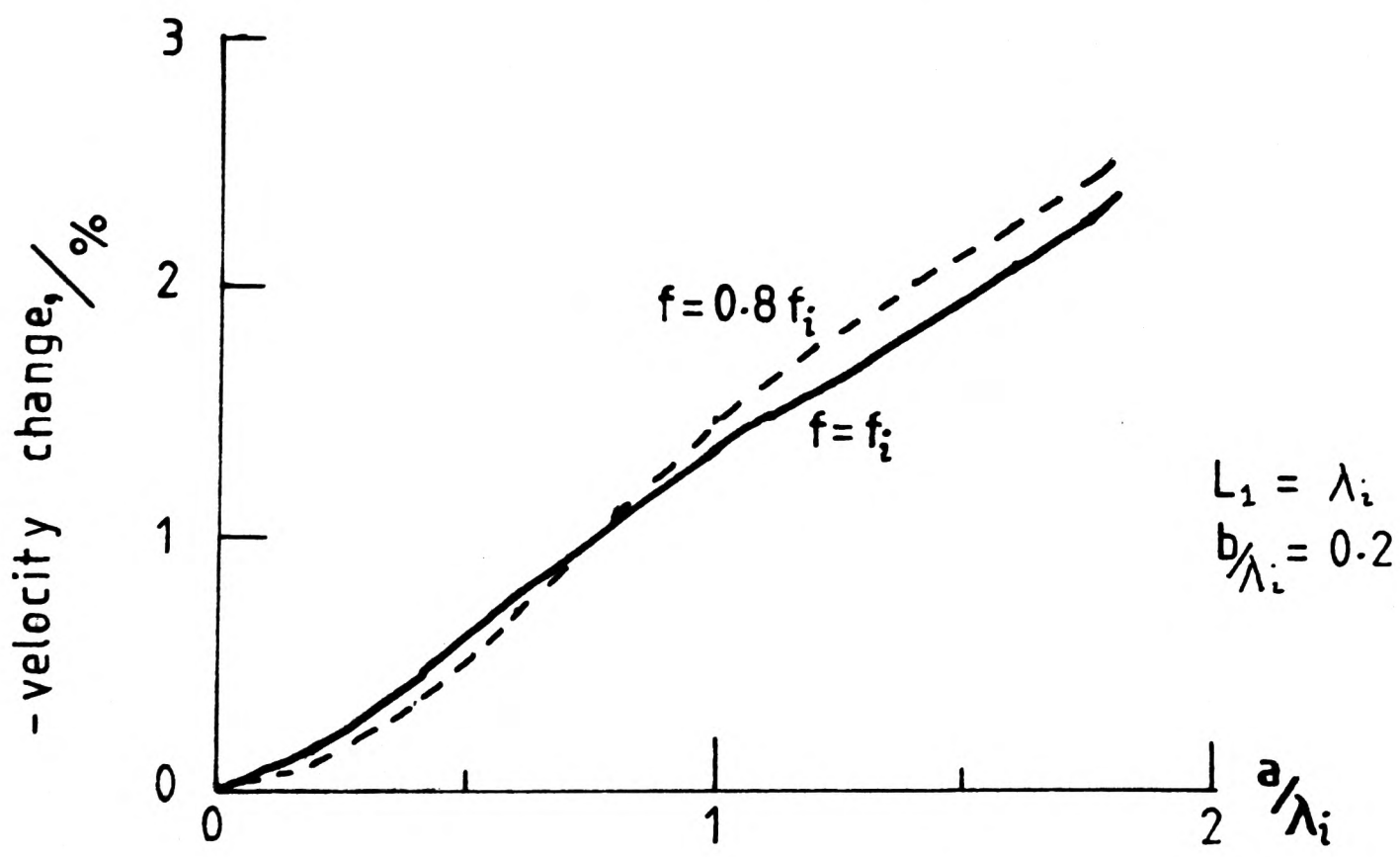


Fig. 5.7: Velocity change in the Z direction due to dots, also normalised to $L_1 = \lambda_i$.

The power lost by a scattering vector is $(r_s N_e)^2 = (r_{os} N_e / (L_1/\lambda_i))^2$, but since the number of scattering vectors is proportional to L_1, total fractional power loss is

$$P = \frac{N_e^2 (r_{os}^2)_{av}}{L/\lambda} \qquad (5.27)$$

The constant of proportionality, found graphically to be approximately 1, applies to the particular array pattern chosen.

Since the material is not isotropic, the angle of inclination of the rows α to the incident beam is not 45° for a reflection through 90°. An error in the angle of inclination causes an error in the amplitude response of the device, although the phase response is, to first order, not affected [5]. The angle is calculated from velocity data. In fig. 5.8, the phase delays of paths **p** and **q** must be equal. It follows that the phase delays of segments C and D must be equal, since the other parts of paths **p** and **q** are identical. An equivalent statement is that C and D represent equal time delays, or

$$\frac{C}{V_{z\,(free)}} = \frac{D}{V_{x\,(array)}} \qquad (5.28)$$

α is given by

$$\tan\alpha = \frac{D}{C} = \frac{V_{x\,(array)}}{V_{z\,(free)}} \qquad (5.29)$$

The angle α thus depends on the **X** direction SAW velocity inside the array and on the **Z** direction SAW velocity in a free surface. The fortunate outcome is that the perturbation of **Z**

velocity by dots is irrelevant to the angle α. Since on $LiNbO_3$, the **X** velocity perturbation is small (about a seventh of the **Z** perturbation), α can be assumed to be constant at 46.8°. (v_z=3488m/s, v_x=3718m/s, is obtained from a best fit of equation (2.2) to experimental data - chapter 2, section 3 - and **X** velocity slowing is calculated from the theory of chapter 3.) The angle is almost identical to the value for groove arrays [47], but this is to be expected from the above considerations.

Finally, information on transducer characteristics is required. The device characteristics are designed by choice of dot array parameters rather than by choice of transducers, so the properties of only one specific transducer (with an appropriate bandwidth) are required. These are obtained most conveniently and accurately by direct measurement. For this purpose, two "devices" each consisting of a pair of transducers facing each other were fabricated, and their characteristics measured using the equipment described in appendices A and B. Tuning inductors were used to reduce the mismatch between the transducers and the 50 ohm source and load, and were adjusted so that the inputs were purely resistive at centre frequency. The characteristics of the two devices were very similar to each other, confirming good repeatability.

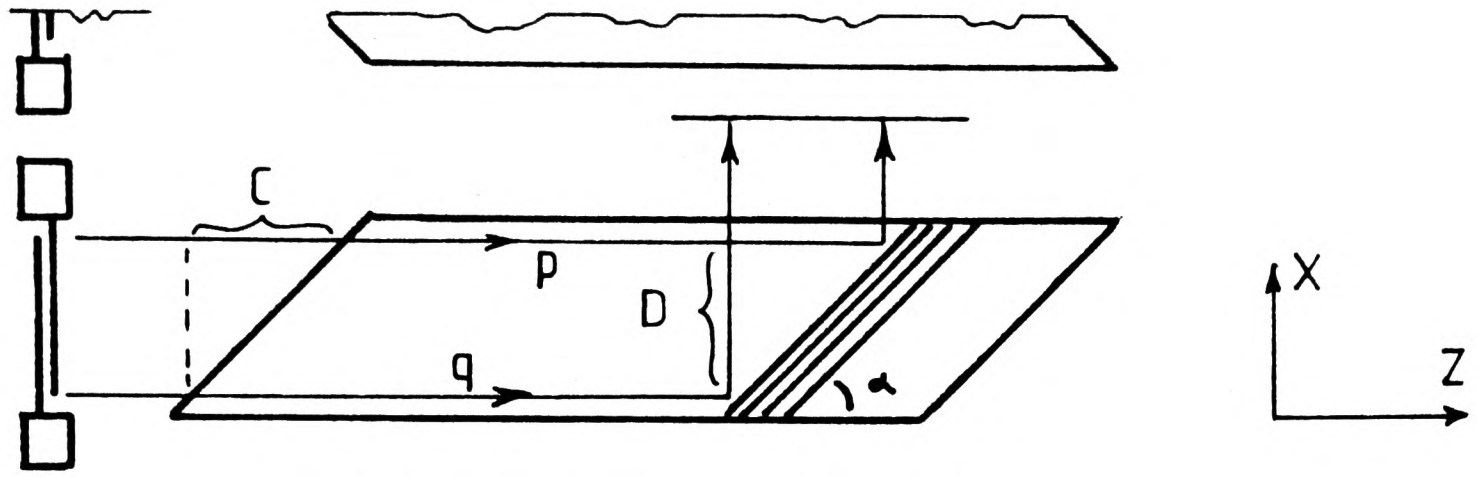


Fig. 5.8: Evaluation of the angle of inclination of the rows

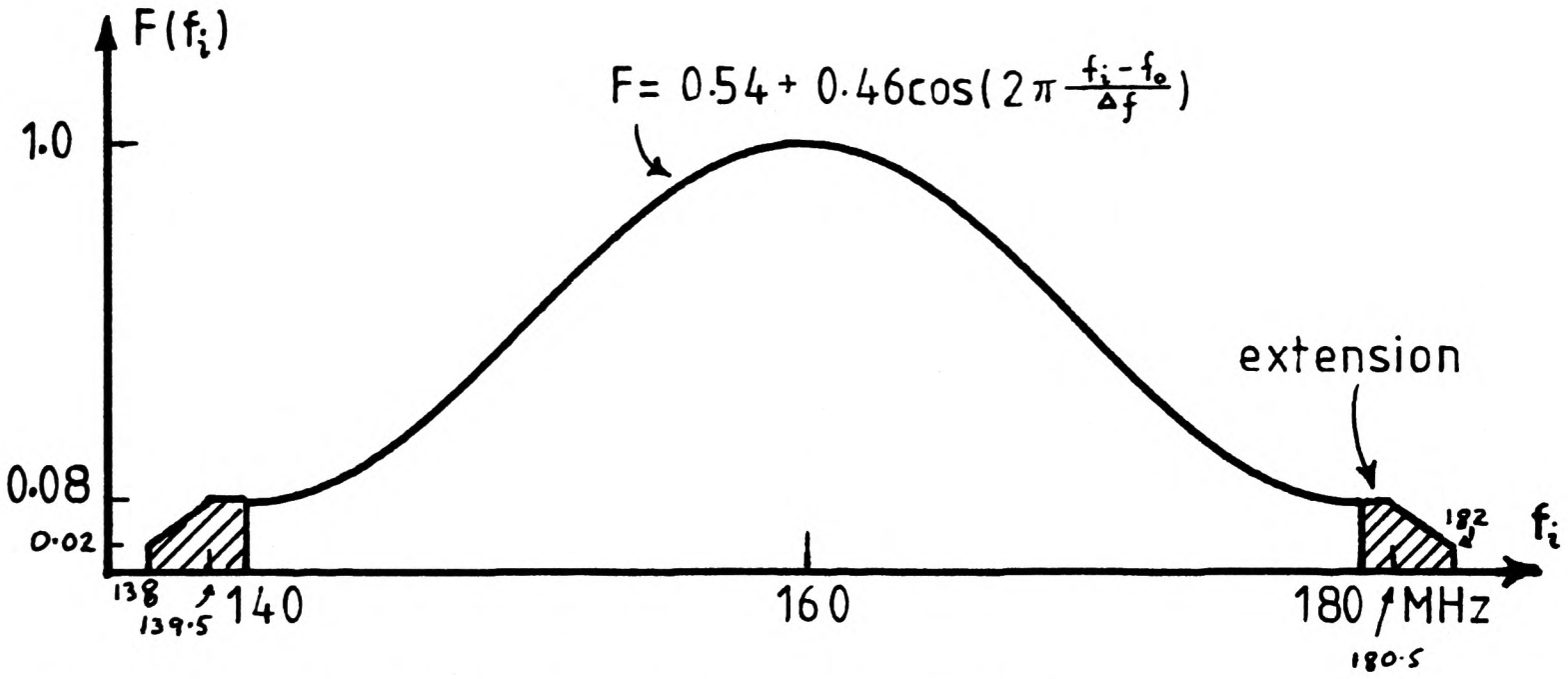


Fig. 5.9: Weighting function used in the RAC.

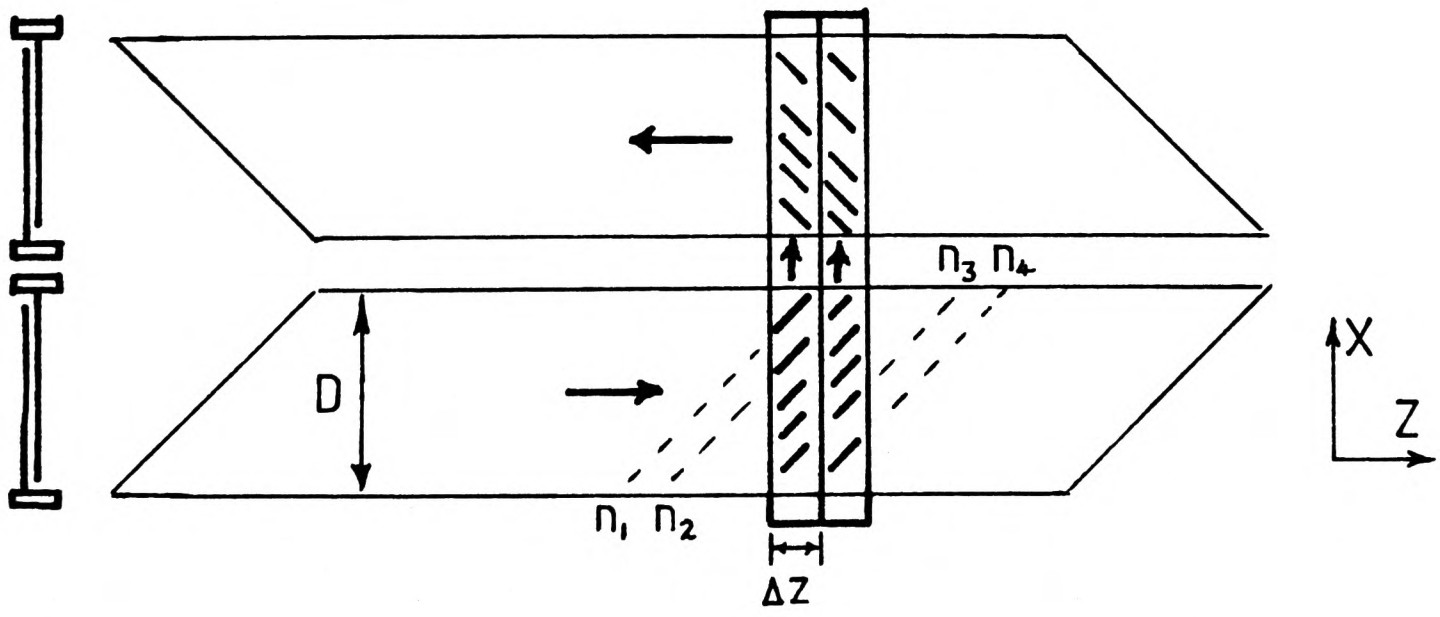


Fig. 5.10: Method of calculating RAC frequency response which is economical in CPU time.

5.7 Computer Calculations for the RAC

The RAC parameters are summarised in table 5.1 . The positions and sizes of 44020 dots evidently have to be calculated by computer. Programs generated to design the RAC are now described. The dot and transducer parameters, described in the previous section, were stored in the program as look-up tables and used in conjuction with simple linear interpolation subroutines. It was convenient to break the job up into a few programs.

The first program finds the value of a/λ_i for each row of dots, b/λ_i being fixed, and the position of the row. The first step is to find the size of the dot in each row required to produce the desired weighting function, (or amplitude frequency reponse). Each row of dots corresponds to a certain instantaneous frequency, given by equation (5.23):

$$f_i = \frac{f_1}{\sqrt{1 - \frac{4B}{f_1^2 T} i}} \qquad (5.30)$$

Because of the principle of stationary phase, and the fact that two reflections are involved, the weighting function is approximately

$$W(f_i) = \sqrt{F(f_i)} \qquad (5.31)$$

where $F(f_i)$ is the Hamming function, modified as illustrated in fig. 5.9. The "extensions" reduce the sudden jumps in weighting near the extremities of the curve. The impulse response $f(t)$ and hence its envelope $a(t)$ are smoothed out even further, since the response at any time t depends on several possible reflection paths from a few rows. The

principle of stationary phase, which assumes a slowly varying **a**(t), becomes a more accurate approximation because of these two considerations. A sudden jump in **a**(t) would cause ripples near the band edges, known as Fresnel ripples and calculated with Fresnel integrals. The exact form of the extensions is arbitrary; no optimisation procedure was thought necessary. The dot "length" **a** is found from $W(f_i)$ after taking propagation loss (from 5.28), transducer amplitude response and scattering loss into account.

Only scattering loss needs any further description. When the size of dot in the **n**th row (whose instantaneous frequency f_n is given by (5.30)) is to be found, the dot sizes for rows 0 to (**n**-1) are already known, so the power loss by scattering can be found. Several localised areas in the array have a grating vector which satisfies the scattering condition, and the size of dot at each of these areas depends on the weighting. For simplicity, the dot sizes at these areas are all assumed to be equal to the dot size at the [illegible] row whose instantaneous frequency is $1.03f_n$; this represents an average size of the dots involved in scattering. The figure 1.03 was estimated graphically by measuring $\underline{\mathbf{k}}_{in}$ ($=\underline{\mathbf{k}}_{out}$) corresponding to the scattering vectors involved, and taking a weighted average.

The scattering and propagation losses each did not exceed 1 dB per array and for half the propagation path, respectively. The response of the transducers was more significant; the frequency response of one transducer was approximately $\sqrt{W(f)}$ or about 7 dB down at the band edges,

leaving about $\sqrt{W(f)}$ of weighting for the array.

The second step is to find the positions of the rows, which determines the phase of the RAC frequency response. In order to reduce computing time, it was convenient to neglect the velocity variation with frequency (which is small) and to apply a correction at a later stage. So assuming that SAW velocity is a function of position only, (5.23) can be used to calculate the positions of successive rows. In that equation, **v** is now a function of dot size and hence of **i**. After the positions z_n of the **n** strips had been found, correction terms were applied for **Z** velocity dispersion (up to 60° per array), **X** velocity slowing (assumed to be a function of dot size only, up to 5° per array) and the transducer phase characteristics, (±8° per transducer), all slowly varying functions. Since each frequency corresponds to a small portion of the array, the correction can be applied independently for each frequency. A few additional remarks on the method of phase correction are made in appendix H. The whole program consumed about 8 seconds of CPU time.

The design program just described assumes the principle of stationary phase. To verify that this was justified, a second program was devised to find the frequency response using the row positions and reflectivities from the first program. The more accurate relation (5.24) was used, except that corrections for propagation loss, reflection loss and **X** velocity perturbation, all very small quantities, were made still assuming stationary phase. These two programs form the basis of an iterative loop to find the required row position

and dot sizes, but the first iteration was found to be sufficiently accurate (within 0.4 dB and 5°) for an experimental device.

It was estimated that a program using (5.24) directly would consume 8 hours of CPU time to find a 400 point frequency response graph (required to resolve Fresnel ripples). The following modification reduced the time to less than 8 minutes. The RAC is divided into small segments as shown in fig. 5.10. Δz is sufficiently small so that the reflected beam can be assumed to be uniform with z in each segment. The frequency response is

$$F_n(\omega) = \frac{\Delta z}{D}\left\{ \text{----} + \left(\sum_{n_1}^{n_3} r_n \exp(-j\phi_n)\right)^2 + \left(\sum_{n_2}^{n_4} r_n \exp(-j\phi_n)\right)^2 + \text{-----}\right\} \quad (5.32)$$

where

$$\phi_n = \omega t_n = \omega \int_0^{z_n} \frac{1}{V(z,\omega)}\, dz \quad (5.32a)$$

The values of $r_n \exp(-j\Phi_n)$ for the 2201 rows are first computed and stored in a look-up table. When computing the summations in (5.32) it is convenient to use the fact that each summation is identical to the previous one except for the addition of one or two new terms, and the removal of a couple of other terms. Proceeding in this way is much faster than adding the seventy to ninety terms in each summation. After computing (5.32), the transducer and compensation terms are included.

The third and final program is very short. Dot sizes are converted to microns, and the positions of the dots along the rows to form the required pattern is found. The program

Table 5.1: RAC Specifications

Centre frequency	160	MHz
Bandwidth	40	MHz
(including "extensions")	44	MHz
Dispersive delay	25	µs
(including "extensions")	27.5	µs
Chirp rate	-1.6	MHz/us
Effective number of synchronous rows	90	at centre frequency
Aperture	80	wavelengths at centre freq.
Transducer: Number of finger-pairs	5 1/2	
Number of rows of dots	2201	per array, including extensions
Number of dots per row	10	
Total number of dots	44 020	
Rows inclined at	46.81	degrees to Z
Columns inclined at	-69.4	degrees to Z
Dot size: a	0.52 to 1.49	wavelengths
(Excluding extensions)	0.66 to 1.49	wavelengths
b	0.2	wavelengths
L1	8	160MHz-wavelengths
L2	1	wavelength

("wavelength"=wavelength at synchronous frequency)

writes the results into a file in GAELIC Language. This file is extended manually to include the transducers and other structures.

CHAPTER 6

PERFORMANCE OF THE PULSE COMPRESSOR

Three nominally identical devices were made from the mask for the RAC pulse compressor designed as described in the previous chapter. The measurements made on the devices to find their performance will now be described, followed by the results and observations.

6.1 Temperature Control of the Devices

Since the devices are made on $LiNbO_3$, they are very temperature sensitive; Z propagating $LiNbO_3$ has a temperature coefficient of delay of about $90x10^{-6}$ per °C [47]. The devices are normally operated in "ovens" (which keep stable temperatures, not high temperatures).

The oven was devised by enlarging the aluminium alloy device package (which houses the device and tuning coils) to include heating elements and a temperature sensor. These additional components were connnected to a commercially available temperature controller, the Electro-Optic Developments' model TC15, which is claimed to maintain a temperature stability of 0.01°C. The package was wrapped in heat insulating material with holes for the electrical

connections. A separate oven/package was used for each device for which temperature control was required.

When the oven is at stable temperature, heat from the coils is required only to make up the loss from its external surface through the insulation. For the temperature to be uniform in all parts of the oven, the heating elements should also be spread over its surface, that is, as near as possible to the sites of heat loss. This reduces the heat flux within the oven, which could have a significant effect even in aluminium alloy because only very small temperature variations are acceptable. Accordingly, the heaters were spread over the outside of the device package using a combination of lumped resistors and 0.18m of 32 SWG Eureka wire, as shown in fig. 6.1 . The main block in the package, on which the device was mounted, had distributed heating by the resistance wire. The lumped, standard resistors were thought sufficient for the lid and bottom because they are not in good thermal contact with the device whose temperature is to be kept constant.

The temperature sensor was a thermistor. Each thermistor used had to be calibrated with an additional resistor, using an accurate thermometer.

The temperature stability of the ovens was not monitored during the measurement of the devices, but showed a drift of less than ±0.02° when tested over a similar duration of time.

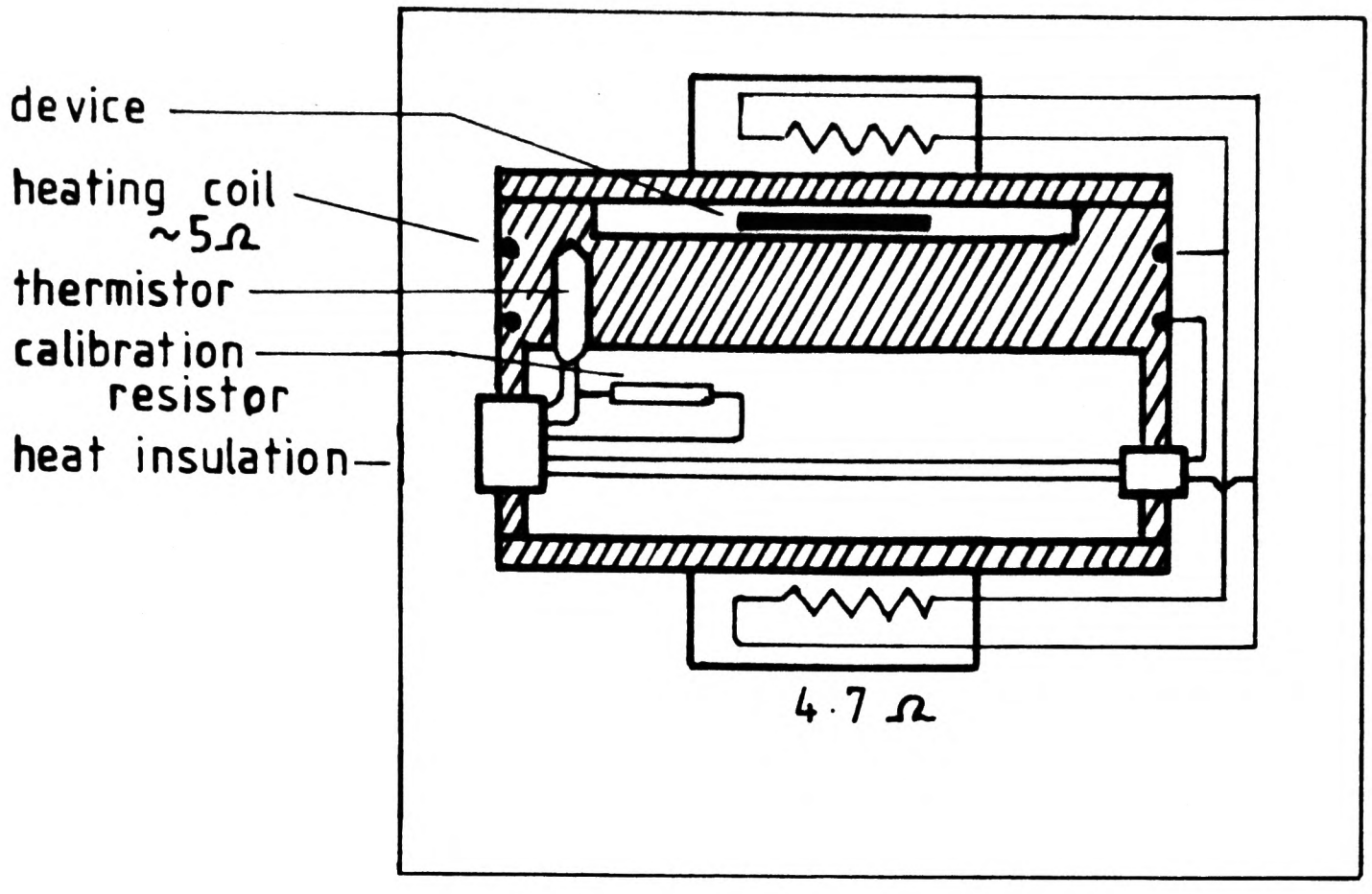


Fig. 6.1: Oven/package for the RACs.

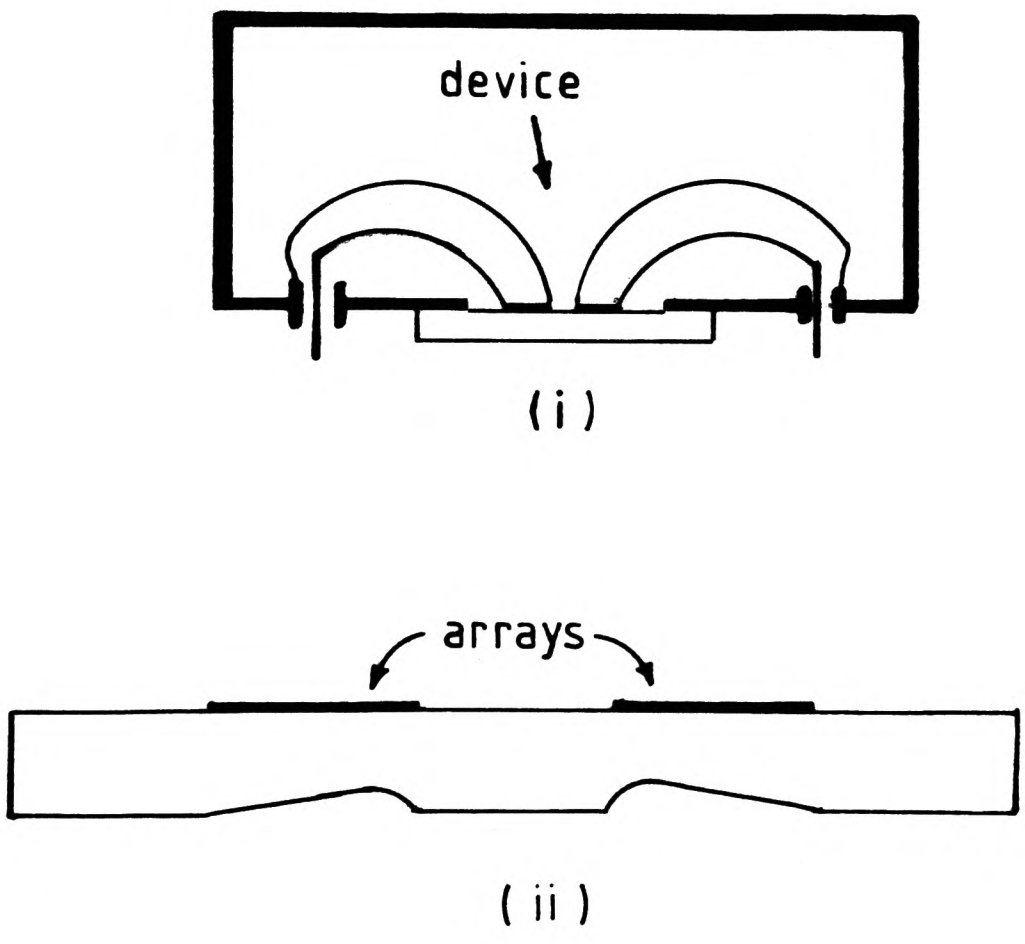


Fig. 6.2: (i) Package for the third device, with the underside of the substrate exposed.
(ii) Shape and position of the grooves.

6.2 Measurement of Frequency Response

Amplitude and phase of the frequency response were measured in separate experiments using measuring circuits and computer software described briefly in Appendices A and B. The conditions under which the measurements were made are summarised here for completeness of record, and are as follows. The input was a pulsed C.W., and the output was time-gated to remove electromagnetic coupling but to accept SAW signals reflected from any part of the array. As with the calibrating transducers mentioned in the previous chapter, the transducer tuning inductors were adjusted so that the combination of transducer and coil in series was purely resistive at centre-frequency. The resistances at this frequency were 45, 45 and 80Ω respectively for the three devices. This indicates variation of metal thickness between the devices about the nominal value of 120nm. (The metal was made thinner than in the dot experiments because the operating frequency is higher.) The possibility of overetching was observed using a travelling microscope giving readings repeatable to ±0.25μm, corresponding to errors in b/λ of 0.01. (The absolute accuracy is probably not as good as this, since 0.25μm is smaller than an optical wavelength.) No overetch could be detected in the first two devices, while the third was found to be overetched by $\Delta(b/\lambda)=0.025$.

The first two devices were measured in ovens kept at 30°C. There were undulations in both the amplitude and phase responses, thought to be caused by bulk waves being generated by the arrays, reflected by the bottom surface of the crystal,

and then mode converted back to SAWs by the array. This will be discussed in the next section. These bulk waves were studied in the third device, which was mounted with its lower surface exposed (fig. 6.2) rather than in an oven, so that the bottom surface could be treated to reduce cohererent bulk wave reflections. Temperature stability was not so important in the third device because the object of study was the undulation of device response with frequency due to bulk waves. Temperature variations would produce similar variations only if it were to rise and fall several times during the half-hour experiment, which is unlikely.

This third device was first measured with its lower surface untreated. The roughness was about 5μm peak-to-peak. Shallow grooves with a maximum depth of about 150μm (about 7 SAW wavelengths) were then abraided into the lower surface (fig. 6.2(ii)) below the arrays, and the frequency response measured again. The grooves were then deepened to about 250μm and the device measured a third time. Finally, rubber solution, which absorbs SAWs and possibly bulk waves, was applied to the lower surface, and a fourth measurement made. Although the 150μm grooves resulted in a dramatic improvement, the further treatment of the lower surface made no significant difference to the frequency response, and will not be discussed again.

6.3 Results and Observations

6.3.1 Amplitude of Frequency Response

The amplitude response of one of the devices (the third, with grooves on the underside) appears in fig. 6.3, together with the design curve. It shows that there is considerable attenuation not accounted for in the design, particualrly at lower frequencies for which the SAW path length is large. The amplitude errors are given in better detail in fig. 6.4, where amplitude deviation from the design curve is plotted for each of the three devices.

The characteristics above about 165MHz are quite consistent from device to device. At lower frequencies, there are undulations in the amplitude error for the first two devices, and for the third before its bottom surface was treated. These undulations present a difficulty because they are not repeatable, and therefore cannot be allowed for in subsequent designs. The undulations disappeared from the third device when the grooves were cut into the lower surface (fig. 6.4), thereby confirming that they were caused by the reflection of bulk waves from the bottom back to the array. Thus with suitable treatment of the bottom of the crystal, it is almost certain that the performance of a device will be predictable.

The bulk wave conversion forms an additional, spurious signal path. Part of the SAW energy is diverted into a bulk wave, is reflected off the bottom of the crystal, and is partially converted by the array back into a SAW whose phase

relative to the original SAW depends on the frequency. Its path thereafter is identical to the path of the required signal. The undulations in the frequency response are thus attributed to constructive and destructive interference. The most probable scattering vectors causing the conversion to bulk waves are those near F in fig. 5.4(iii). Each scattering vector can reflect a band of frequencies, because a band of bulk wave vectors $\underline{\mathbf{K}}_b$ can have the same surface component $\underline{\mathbf{k}}_b$ which satisfies the scattering condition. It will be recalled that it was possible to position the scattering vectors by choice of array pattern to avoid scattering by LSWs and bulk waves, except for the scattering vectors near F. The additional path length of the spurious signal depends on the thickness of the crystal, which is not consistent between crystals to a fraction of a wavelength. This explains the lack of repeatability of the devices. The grooves cause the bulk waves to be reflected in a different direction, so that, on returning to the array, they cannot be converted back to SAWs by any of the scattering vectors.

The possibility that the array picks up bulk waves generated by the transducer can be ruled out. The wave vector of any bulk wave generated by the transducer has a surface component $\underline{\mathbf{k}}_b$ approximately parallel to **Z**, and is always smaller than $\underline{\mathbf{k}}_{saw}$. By examining the scattering vectors of the array (fig. 5.4(iii)), it can be seen that no scattering vector satisfies the scattering conditions (3.28,3.29).

After the undulations in the response are removed, there remains a repeatable loss corresponding to attenuation which is approximately proportional to path length in the array. Low frequencies, reflected by the far end of the array, are attenuated more than the high frequencies (fig. 6.3). Since it is already known that bulk wave generation is significant, this attenuation is also attributed to the generation of bulk waves. Thus the SAW is affected both when the bulk wave is generated and later when it is reflected back to the array. The slowly varying part of the error is attributed to the generation, while the undulating component has been shown to be a result of the reflection.

The effect of bulk wave generation can be observed in the third device, in which the other source of error (bulk wave reflections) have been removed. The almost straight line (fig. 6.4) indicates that attenuation is proportional to distance travelled by the SAW in the array despite the fact that dot length varies according to the weighting, over a factor of nearly two. This suggests that the attenuation is broadband and is independent of weighting. If this is true, allowing for the attenuation in future devices will be particularly easy, even in new designs with different specifications. When the dot lengths are increased to increase SAW reflectivity, the attenuation is not affected further, so iterations would converge rapidly on the desired result. Since two reflections are involved, reflectivity of the low frequency portion of the array has to be increased by only 60%. Alternatively, by changing reflectivity in the range ±25%, the **relative** weighting can be corrected. It

should be noted that it is accepted practice to re-design a device after a first iteration, as suggested here.

It was confirmed that this loss was not due to beam steering on a misaligned crystal. This could have caused a similar error. Because of the anisotropy of $LiNbO_3$, the direction of power flow is not parallel to the wave vector if the wave vector is not exactly along **Z**. Thus a fabrication error of this kind can result in the arrays being only partially illuminated, with some of the energy "missing" the reflecting rows of dots. The loss would be greatest at the far-end of the array, so low frequencies would be affected most; beam steering could have been a possible explanation for the attenuation. However, the direction of the beam was ascertained by a potential probe, and no measurable beam steering was observed. (The probe could be positioned to an accuracy of better than 1/20 of the beam aperture.)

Device errors are summarised in table 6.1 .

6.3.2 Phase of Frequency Response

Equation (5.25) shows that the phase response should be quadratic. The actual phase response of a down chirp device can be expressed as

$$\phi = \phi_o + \alpha f + \pi \frac{T}{B} f^2 + e(f) \qquad (6.1)$$

where **e**(f) is the phase error. It is conventional to calculate Φ_o, α and $\pi T/B$ for a least-square-error fit, that is for a minimum RMS value of **e**(f). Device performance can then be presented as the error in $\pi T/B$ and a plot of **e**(f). The errors

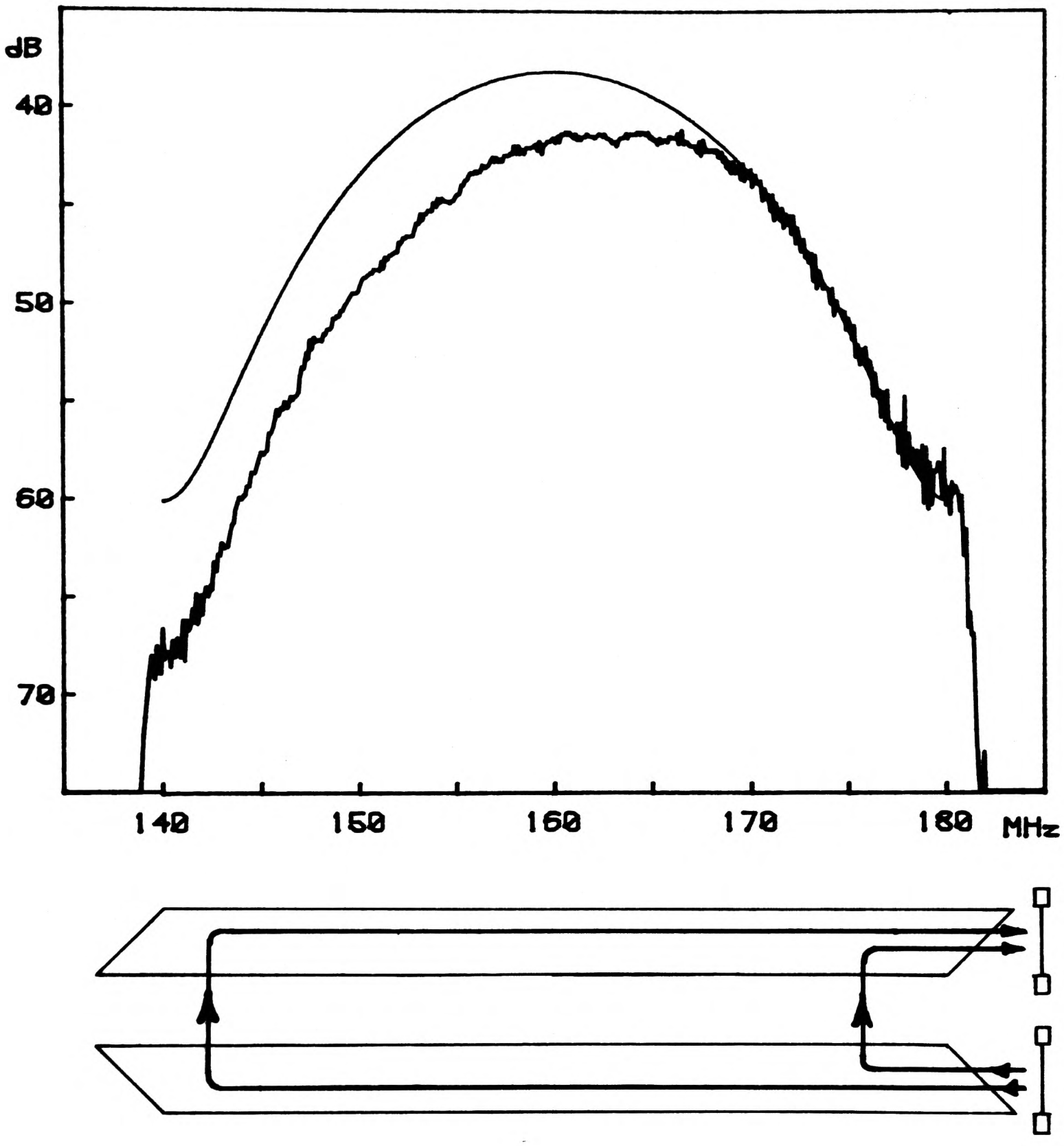
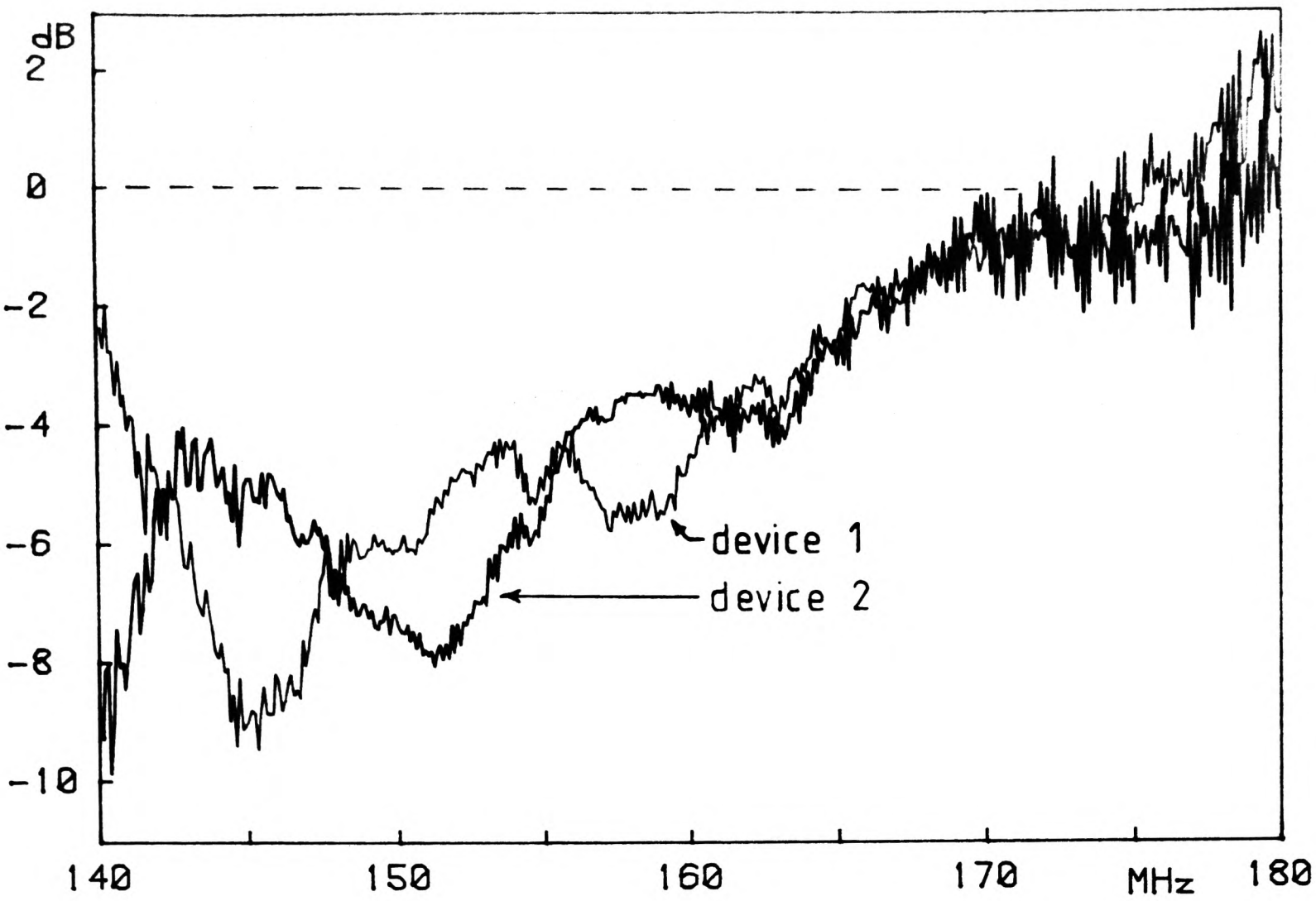


Fig. 6.3: Amplitude of frequency response of one of the RACs (lower line), shown with the design curve.



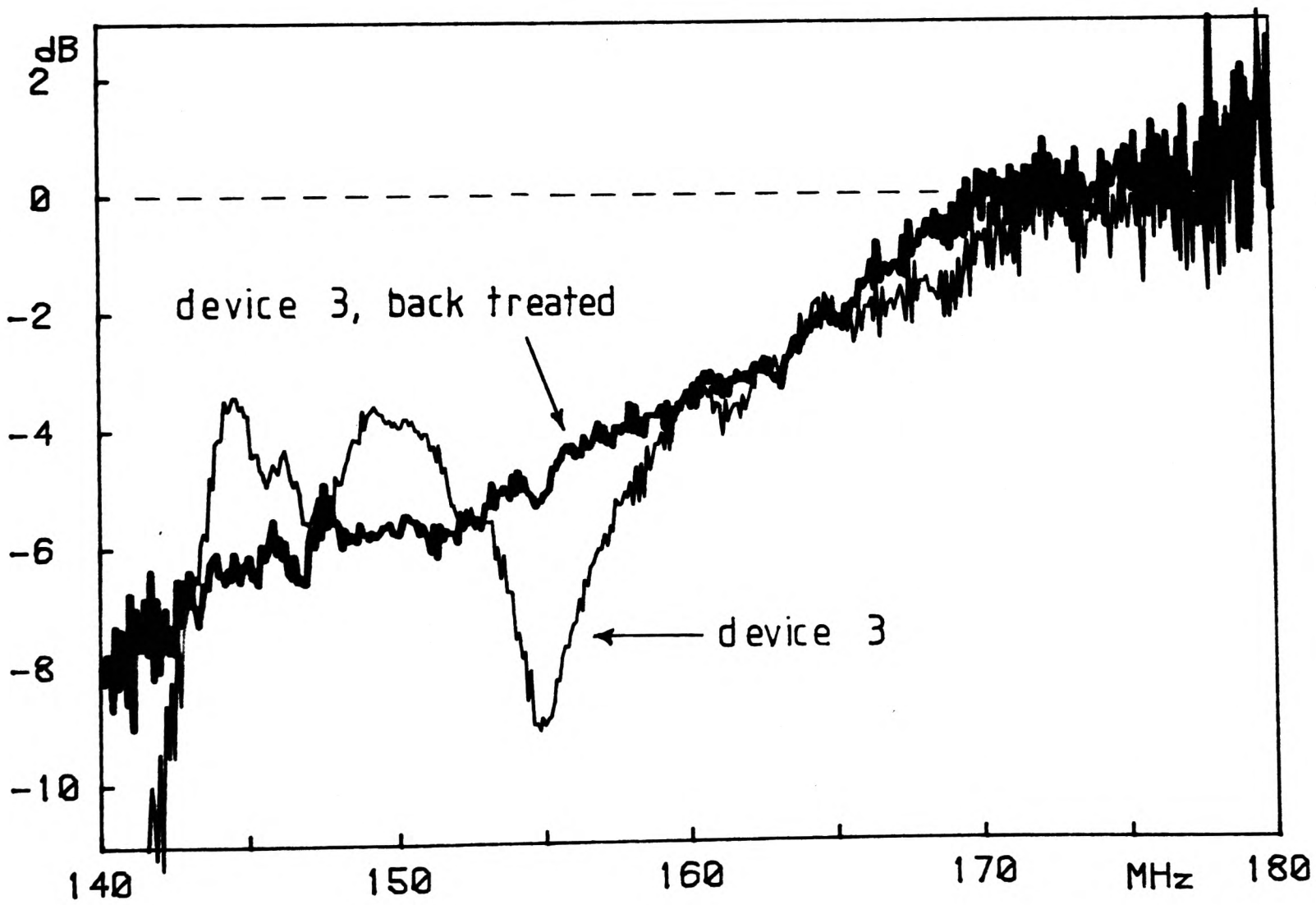
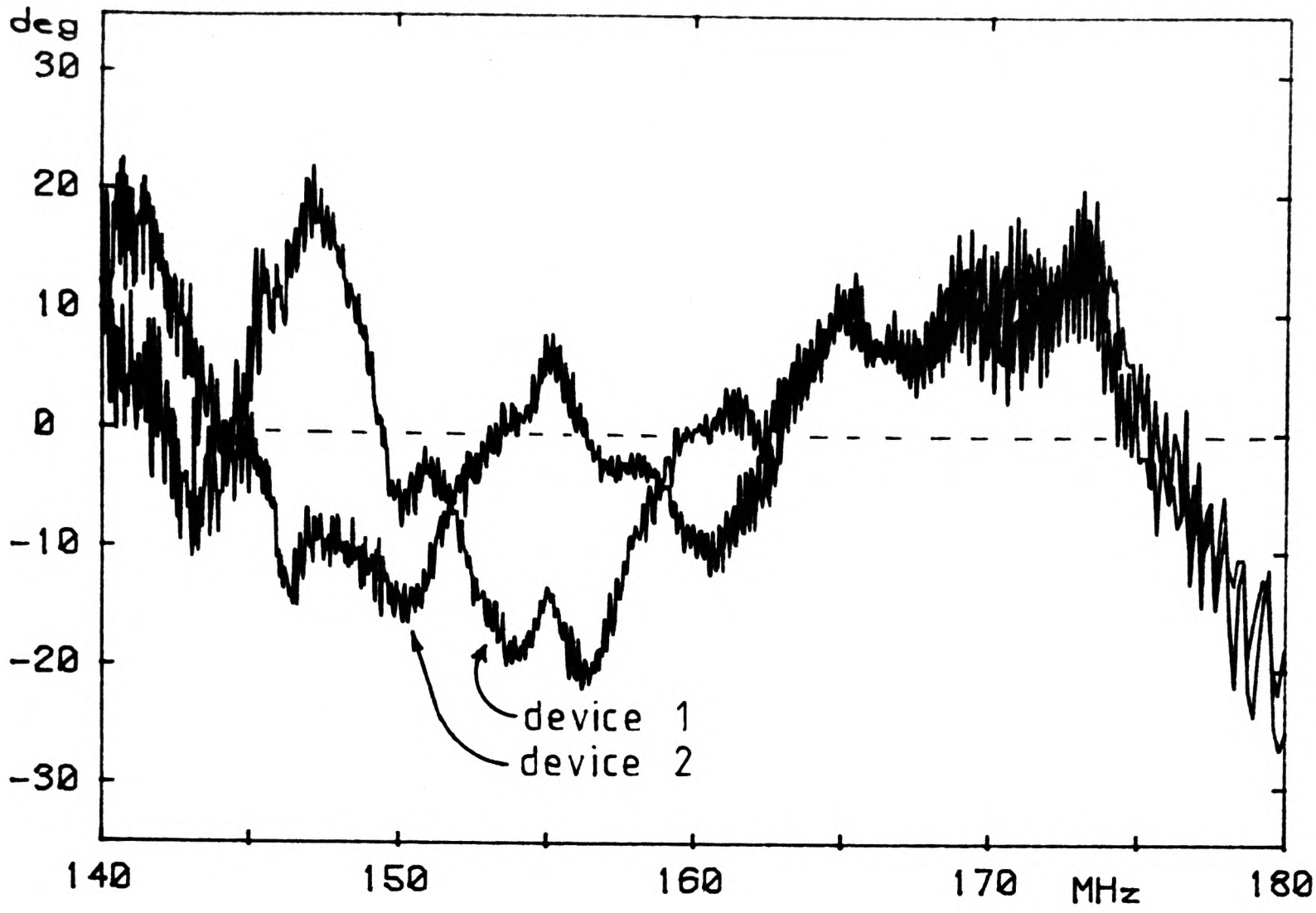


Fig. 6.4: Amplitude error of the RACs.



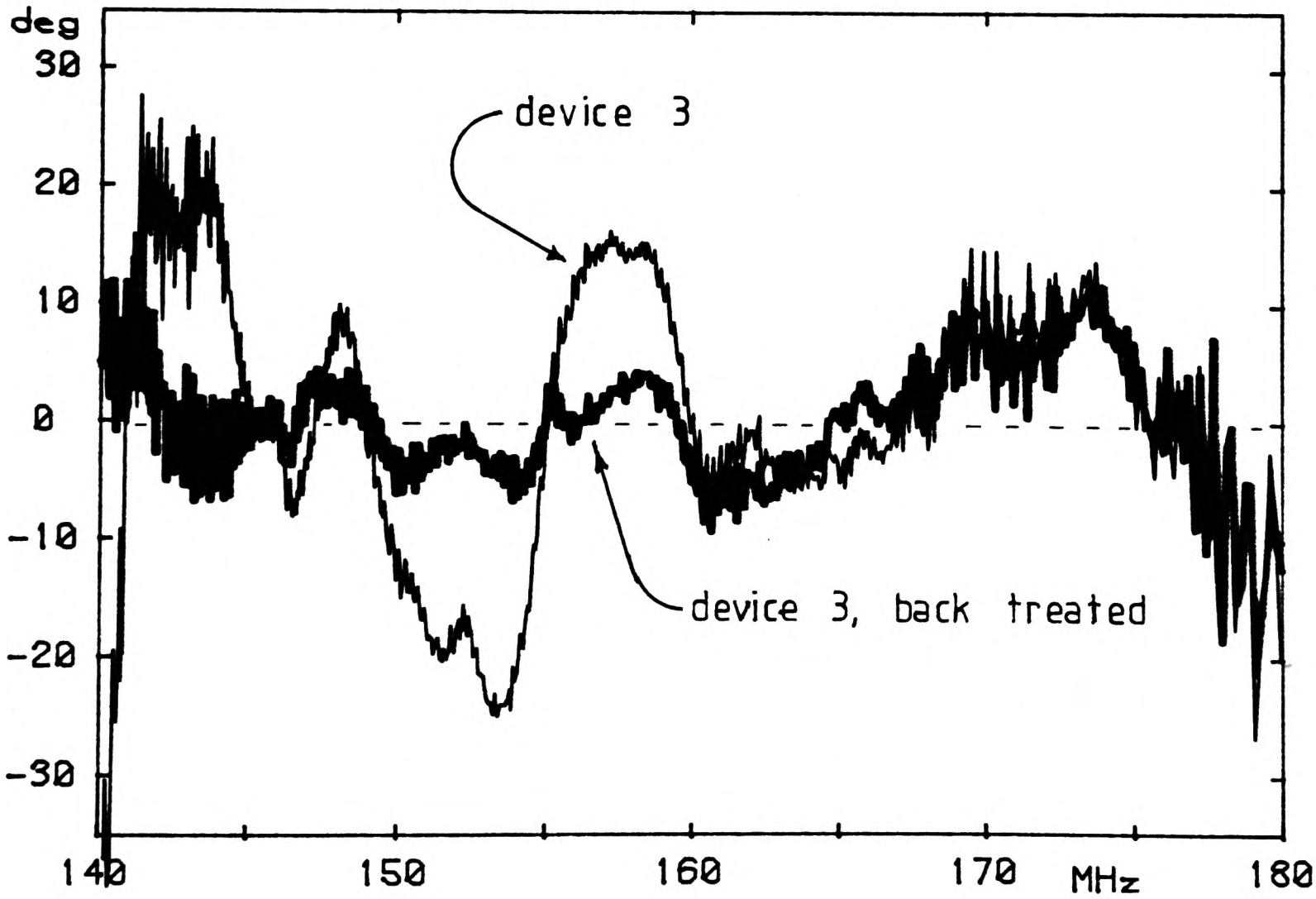


Fig. 6.5: Deviation from quadratic phase of the RACs.

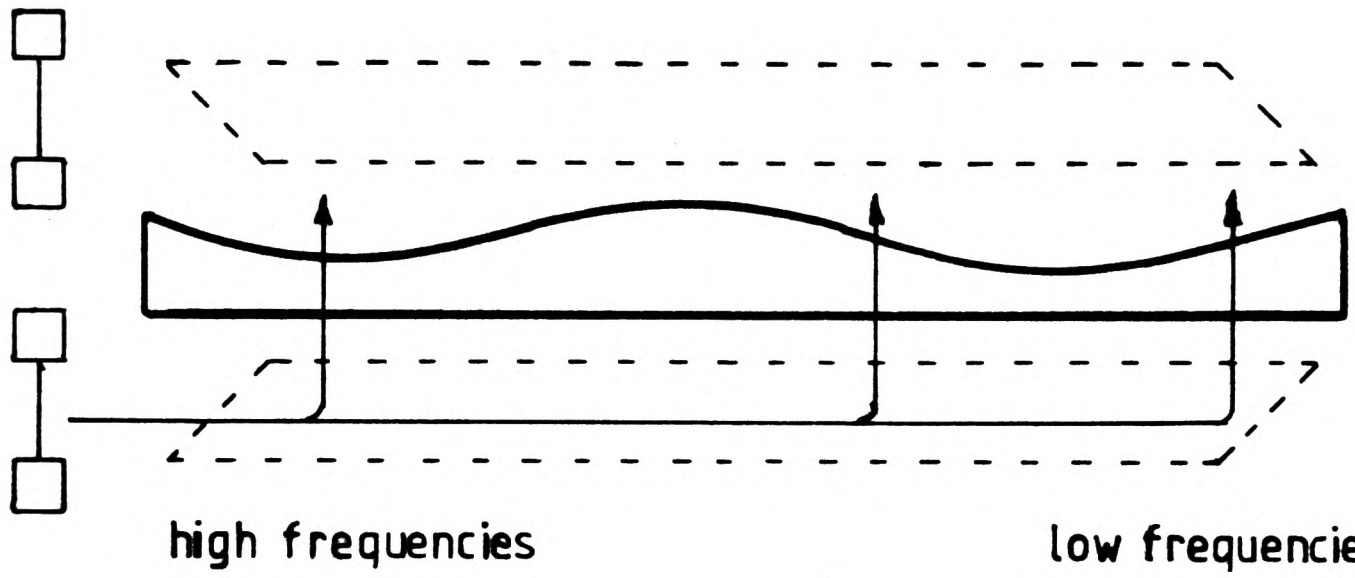


Fig. 6.6: The phase plate.

in Φ_o and α are usually irrelevant in an experimental device, since the corresponding group delays are respectively zero and constant.

Plots of **e**(f) are given in fig. 6.5. With the lower sides of the substrate untreated, the RMS phase errors of the devices are less than 12°, which is entirely satisfactory. However, RMS error was reduced to only 6° in the third device by the grooves in the bottom surface of the substrate. As with the amplitude error, phase error is reasonably repeatable from device to device above about 165 MHz, but not below that frequency, where unpredictable undulations occur. They are substantially reduced in the third device by the grooves in the bottom surface (fig. 6.5). It is clear that much of the phase error and much of its uncertainty is due to bulk wave reflection off the bottom, which is relatively easy to prevent.

The errors in (**T**/**B**) of the devices are recorded in table 6.1 . The first two devices have errors consistent with each other. The third device is different, because it was not operated in an oven for reasons already discussed.

It is common practice to add a phase plate between the reflecting arrays to adjust phase response after each device has been fabricated. The phase plate is a metal film (fig. 6.6) which reduces the velocity of the SAW, therby modifying the phase. Since the frequencies are spatially separated by the RAC, any arbitary error function **e**(f) can be accomodated provided that it is slowly-varying.

Table 6.1: Summary of RAC Errors

	Device 1	Device 2	Device 3 (no oven) Bottom untreated	Device 3 (no oven) Bottom treated	
AMPLITUDE Mean amplitude error	-3.5	-3.6	-4.6	-3.2	dB
PHASE Error in quadratic term (T/B) (designed: 0.625)	0.0019	0.0017	0.0012	0.0010	µs/MHz
RMS deviation from best-fit quadratic	11	10	12	6	deg

6.4 Summary of Observations

1. The non-repeatable part of the amplitude error, which varies rapidly with frequency, is largely due to bulk waves reflected from the bottom of the crystal being converted back into SAWs, and is easily reduced substantially.

2. The slowly varying part of the amplitude error is repeatable, and can therefore be compensated for in future designs. It is approximately proportional to the distance travelled by a SAW of a particular frequency inside the arrays, and probably independent of frequency and dot size for the range of sizes and frequencies used in the RAC.

3. The phase error can be as low as 6° RMS if the bottom surface of the substrate is suitably treated. The error in chirp rate is probably predictable and can be compensated for in future devices. Both these errors can also be corrected using phase plates designed for each individual device.

CHAPTER 7

CONCLUSIONS

7.1 The Outcome of this Project

The aim of this project was to study the reflection of SAWs on piezoelectric materials by thin metal dots, for the purpose of making devices. The investigation has been both theoretical and experimental.

The work on the theory has produced a method by which the relevant dot properties, reflectivity, scatter and velocity can be predicted. The model is numerical, although a simplified, approximate analytical version was also produced.

The theory has been verified for aluminium dots on $LiNbO_3$, a material commonly used in SAW devices. Reflectivity was measured for a wide range of dot sizes. Velocity change and scatter by a dot were also investigated. Although only one material was used, it is believed that the theory can be applied, with minimal additional experments, on other strongly piezoelectric materials where electrical effects are the primary source of reflection.

The essential characteristics of dots can now be predicted. Reflectivity can be predictably varied over a wide range by selecting appropriate dot sizes. Dots with particular useful properties, such as minimal sensitivity to overetch, or deliberately low or high reflectivity, can be identified, so the size of dot best suited to each application can be used. Similarly, if there is a need for dots which reflect preferentially in a particular direction, the appropriate shape can be sought from the theory. A further advantage of a detailed knowledge of dot behaviour is that the number of design iterations is reduced. This will probably be the main use of the velocity data: the variations of velocity with frequency can be allowed for in the initial design.

The theory was tested by using it to design a device. The RAC produced behaved as predicted except that the SAW was attenuated in the reflecting arrays, presumably because of power loss to bulk waves generated. It is expected that this can be easily corrected in future designs, since the loss (for a range of dot sizes) can be estimated from the RAC. The other problem, the undulating error caused by bulk wave reflections, has already been solved by treating the back surface of the substrate.

It is almost certain that arrays of thin dots are suitable for SAW devices, and that their behaviour has now been characterised sufficiently to design them. It remains only to test this view by designing more devices.

7.2 Possibilities for Further Development

Further research could be done both on specific devices and on the general behaviour of dots.

Although the array pattern of the RAC (chapter 5) was chosen so that power loss and spurious signals due to conversion to LSWs and bulk waves would be small, no attempt was made to produce an algorithm to find the ideal array pattern for this purpose. This may well be useful, since it was discovered that conversions to bulk waves are probably a significant cause of attenuation. In order to do this, it would first be necessary to have a method of predicting the level of LSW and bulk wave generation. Further experiments could be done, and an additional theory developed. It may be that the present theory on SAW to SAW reflections, with minor changes, will be sufficient for these additional questions.

A model to determine SAW to LSW reflections could be useful for another reason. It is known that a wideband up-chirp RAC using metal strips would be difficult to design because of LSW generation [21]. Too much of the incident power would be reflected into the LSW before the "wanted" reflection further along the array. This problem is also present with dot arrays. (It has been seen in chapter 5 that the opposite is true of down-chirp RACs, where the wanted reflection takes place before the SAW-LSW reflection, so that the latter removes power only from a beam of waste energy.) There is a possibility that a dot size exists in which the SAW-LSW reflectivity is zero, while the required SAW-SAW reflection is finite. This dot size would be useful for the

up-chirp RACs provided that the level of bulk wave generation is not too high. LSWs might also be supressed by grouping the dots such that the LSWs from the dots in each group cancel.

The final topic to be proposed here concerns the use of "hole" arrays instead of dots. Although the reflectivity of holes is reported in chapter 4, it will be recalled that work on the subject was discontinued to concentrate on dots. In addition to the possible advantages mentioned in chapter 4, the hole arrays may also be better than dots in in-line RACs (ILRACs), in which the reflectors reflect the SAW through 180°, and the incident and reflected beams are separated with a multi-strip coupler [48]. The difficulty with this configuration is that it is not possible to use a phase plate to correct the phase response of a device.

A substitute for the phase plate can be made from a hole array consisting of two layers of different metals. The idea of a double metal layer has already been used for another purpose [13]. Part of the top layer can be selectively etched away, leaving the lower layer intact. Its proposed use as a phase plate substitute is illustrated in fig. 7.1 . Since the reflection mechanism can be chosen to be primarily electrical, the partial removal of one metal layer does not significantly affect the amplitude of the reflection. However, there is a slight difference in SAW velocity between the singly and doubly-plated areas due to mechanical effects. Since the phase correction required is typically a few degrees over a few thousand wavelengths, only a very small velocity difference is required. The phase lag of the nth row is the

average phase lag of all the dots in that row, each of which is determined by the distance traversed by the SAW through singly-plated and doubly-plated areas. The phase shift of the nth row is therefore proportional to the area of metal removed between the nth row and the transducer. This can be adjusted by changing D_n, the width of the singly plated region at the nth row. The widths higher up the array, D_i, cannot be used to make this adjustment since they have already been fixed to determine the phase lag of other rows.

The holes could be positioned to aim for the pattern in fig. 7(ii). Both positive and negative phase errors can then be accomodated by removing more or less metal than this nominal pattern.

It is not known whether the boundary between the single and double layers can be a smooth curve instead of the step function shown. The possible difficulty is that the SAW would be incident on the smooth curve at a very small glancing angle, which could result in a large reflection even if the mismatch between the regions is very small.

Although a significant advance has been made on the subject of dot arrays, it is clear that there is room for further development. It is however hoped that the work already done will be useful for a wide range of SAW devices.

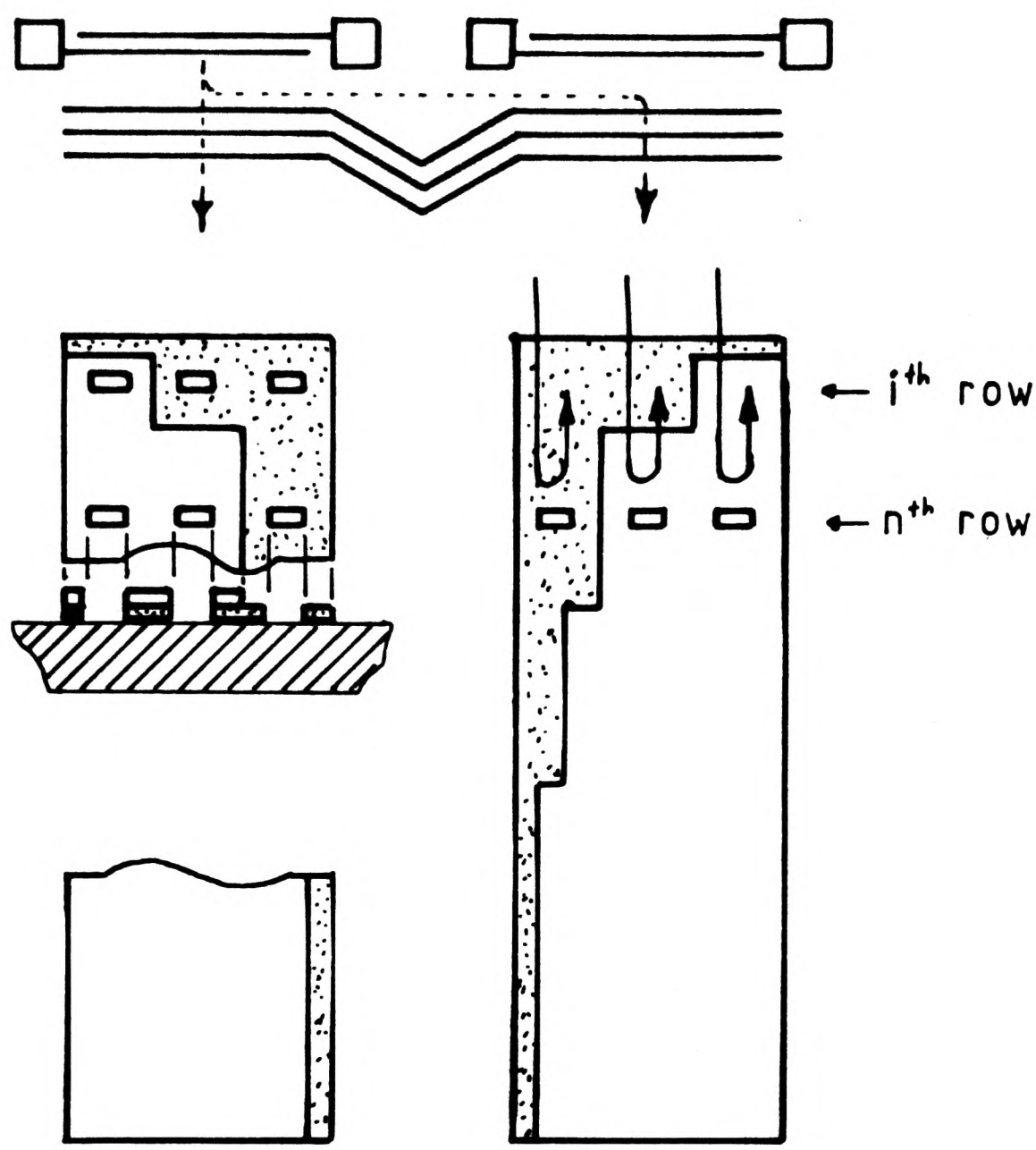


Fig. 7.1: (i) ILRAC with subtitute phase plate, involving a bi-metal layer.

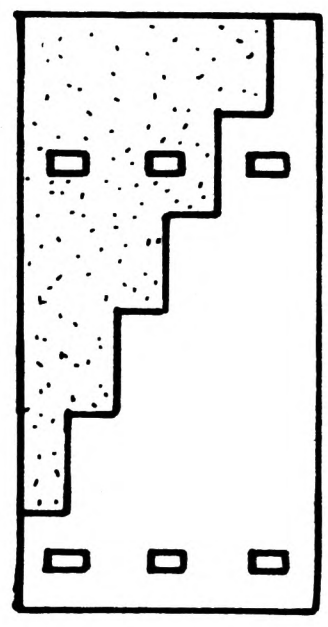


Fig. 7.1: (ii) Nominal shape of top metal layer to which device is designed, permitting both positive and negative corrections.

Appendix A

The Insertion-loss Measurement System

The insertion-loss system was implemented by A. Stove, based on a similar but slower system at RSRE (Royal Signals and Radar Establishment), and modified by R. C. Woods.

The response of a typical device to be studied in this measuring system to a burst of C.W. is given in Fig. A1. Output burst B is due to electromagnetic coupling between input and output transducer leads. Pulse A is the required output due to the acoustic wave. By performing a sample-and-hold at time T, only the wanted signal is obtained.

Even though an r.f. pulse is narrow-band rather than single-frequency, an accurate frequency response is obtained. This is easily verified in the time domain. Let the impulse response of a device be $f(t)=f_B(t)+f_A(t)$ where $f_B(t)$ is due to e.m. breakthrough and $f_A(t)$ is due the acoustic wave. The output $V_A(T)$ at a sampling time T due to the acoustic wave only is related to the input by

$$V_A(T) = \int_{-\infty}^{\infty} f_A(t') V_{in}(T-t') \, dt' \qquad (A-1)$$

$V_{in}(t)$ is a constant amplitude r.f. pulse. $f_A(t')$ is non-zero only for a range of t', $T_1 < t' < T_2$, determined by the possible acoustic paths in the device.

$V_{in}(T-t')$ can thus be zero outside the corresponding range $(T-T_2) < (T-t') < (T-T_1)$ without affecting the integral. This choice of V_{in} makes

$$V_B(T) = \int_{-\infty}^{\infty} f_B(t')\, V_{in}(T-t')\, dt' = 0 \qquad \text{(A-2)}$$

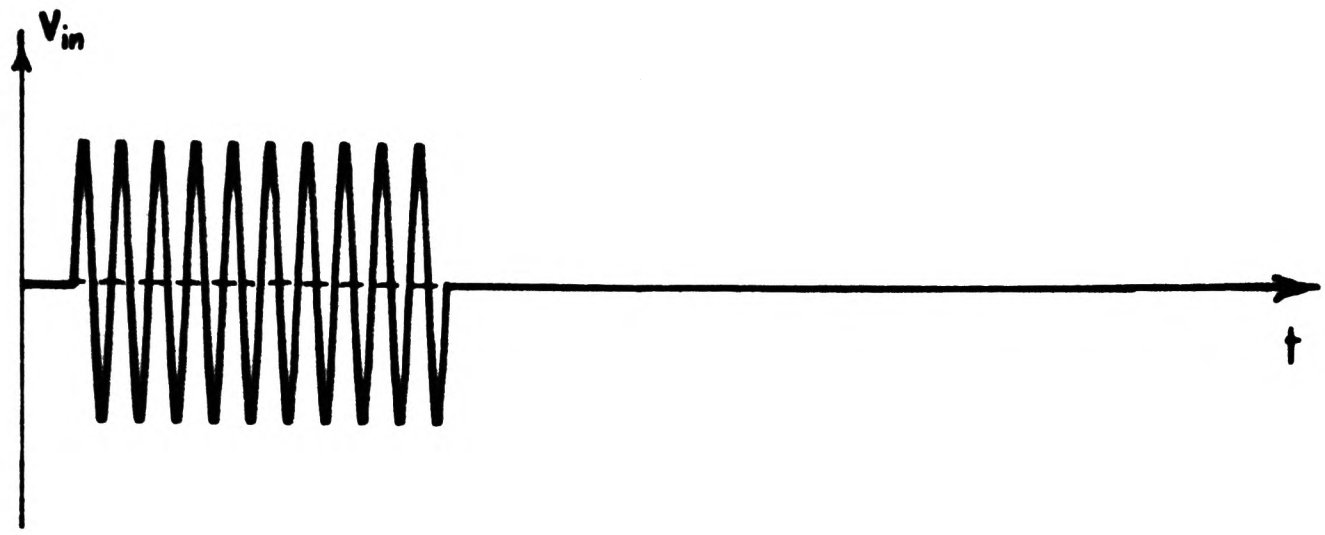
if $V_{in}(T-t')$ is zero whenever $f_B(t')$ is non-zero.

This system is desirable for experimental devices for two reasons. Firstly, the devices (which are not mass-produced) do not have to be carefully packaged to reduce B. Secondly, the e.m. breakthrough is usually significant only at those frequencies ω where the response $|F(\omega)|$ is small. Even if the properties of a device at small values of $|F(\omega)|$ cannot be used directly, they can aid in the understanding of experimental devices.

A block diagram of the measuring system appears in fig. A2. All the equipment is controlled automatically by a Hewlett Packard 9825 Programmable Calculator, either directly or through a Hewlett Packard 6940B Multiprogrammer (which contains an analogue to digital converter and other interface units), except that timing and duration of pulses from the Pulse Generator are pre-set manually. The Pulse Generator controls the sample-and-hold circuit (the Linear Gate) and also switches the input on and off via the Mixers.

10dB Attenuators at the input and output ports of the device ensure that the source and load seen by the SAW device are 50 Ω. The signal is switched alternately through the device and through Programmable Attenuators, and the Attenuators adjusted automatically until the output level

through them is as near as possible to that through the device. 1dB is the smallest step in the attenuators, so fractions of a dB are then found by comparing outputs. The signal level of the Hewlett Packard 8660C Frequency Synthesiser (signal generator) is controlled so that this comparison is done as close as possible to 9mV, near the maximum output for which the Detector does not overload. At very high insertion losses, the comparison is executed at a lower output since the input power amplifier reaches its maximum power.



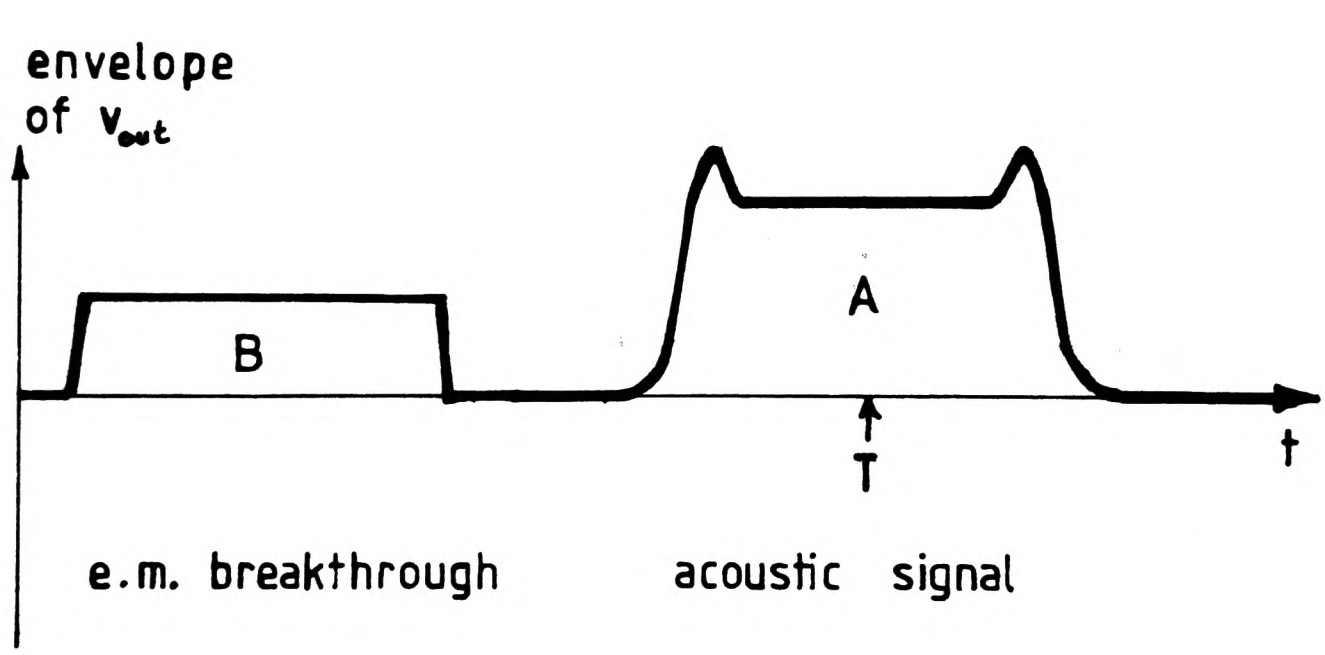
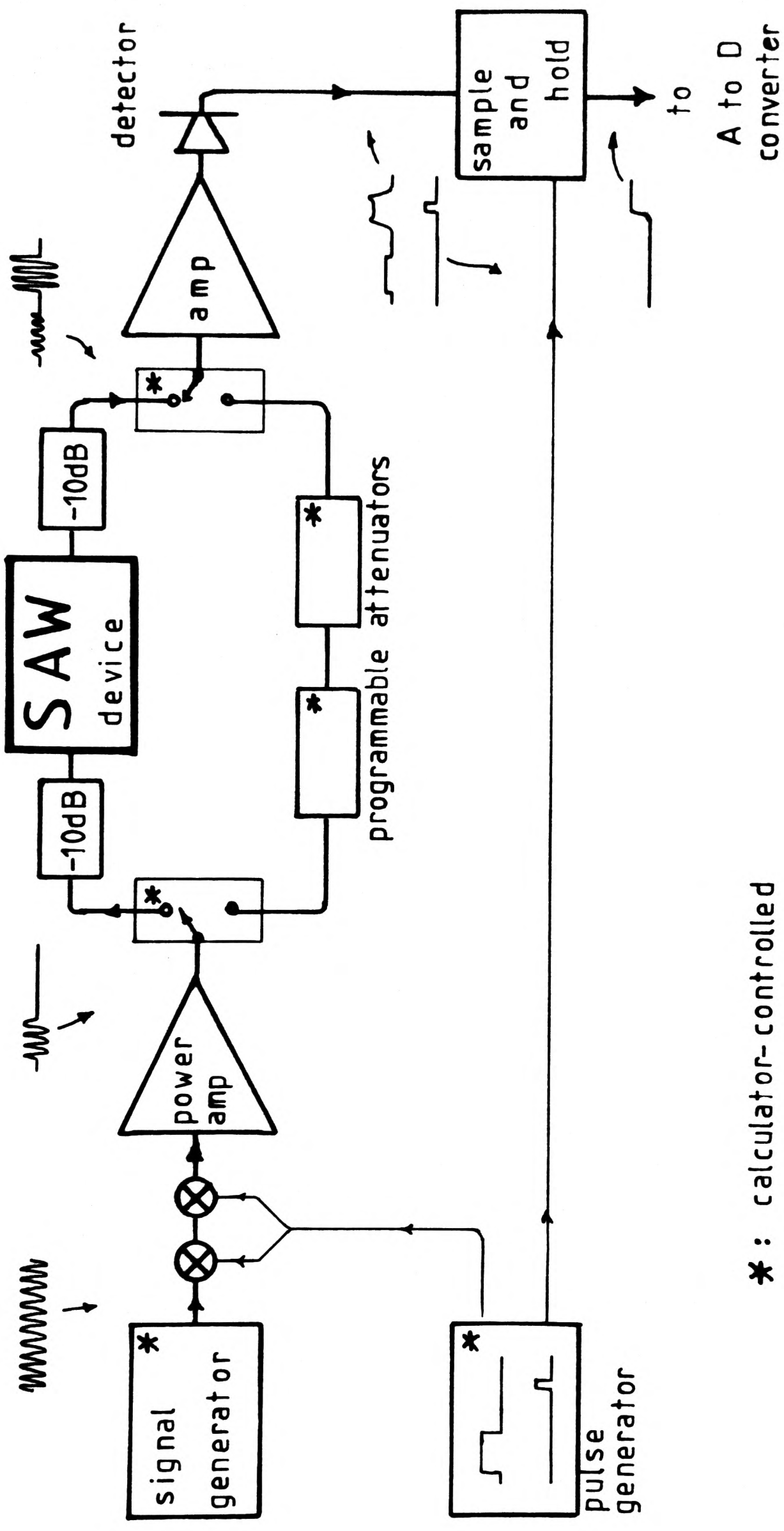


Fig. A1: Input to and envelope of output from a SAW device which has not been well packaged; output shows pulses due to both electromagnetic breakthrough and the acoustic signal.



*: calculator-controlled

Fig. A2: Simplified block diagram of the insertion loss measuring system.

Appendix B

The Equipment for Measuring Phase Response of a RAC

The software and circuit for measuring phase response was developed by R. C. Woods. All calculations described here are performed automatically by the calculator program which also controls the experiment. Equation (5.17) indicates that the phase response of an ideal RAC varies quadratically with frequency, with the form

$$\phi(\omega) = \phi_o + \alpha\omega + \beta\omega^2 \qquad \text{(B-1)}$$

Since $\Phi(\omega)$ varies by several hundred cycles over the passband of a typical RAC, sufficient information on $\Phi(\omega)$ can be obtained if all the frequencies at which $\Phi(\omega)$ is an integer number of cycles is found. In other words, the values of ω_n such that

$$\phi(\omega_n) = 2\pi n \qquad \text{(B-2)}$$

are to be determined.

A block diagram of the measuring circuit appears in fig. B1. The mixer A chops to c.w. input into pulses, so that direct electromagnetic coupling can be distinguished from the SAW signal, as described in Appendix A. The amplifiers are assumed to give negligible phase variations over the frequency band measured. The attenuators ensure that the source and load impedances as seen by the SAW device are approximately 50 ohms. The phase sensitive detector is a mixer. As frequency is varied, the mixer output is maximum when the reference and the output from the SAW device are in

phase.

Since reading the output at each frequency consumes a finite amount of time, it is desirable to find the maximum output using only a small number of readings. A series of readings is taken at frequencies ω_{n1}, ω_{n2}, ... separated by $\Delta\omega$ is taken until a maximum point is passed (fig. B2). The curve near the maximum point is then approximated by a parabola passing through the last three points, and the frequency ω_n corresponding to the maximum point on the parabola is then calculated analytically. This gives ω_n to considerably better resolution than $\Delta\omega$.

Two options on the presentation of the data are available in the software. In the first, Φ_o, α and β in (B-1) are determined to give a least-mean-square deviation from the measured data. (B-1) with Φ_o, α and β so determined then represents the quadratic curve with the best fit to the data. The deviation of the measured phase from this best-fit curve is then plotted as a function of frequency, and the RMS value of the deviation is given. The coefficients α and β are also given by the program. Φ_o is usually irrelevant, as it corresponds to zero group delay.

In the second option, the designed value of β is entered into the calculator. Φ_o and α are chosen for the best fit curve, and the deviation of the measured data from this curve is then plotted.

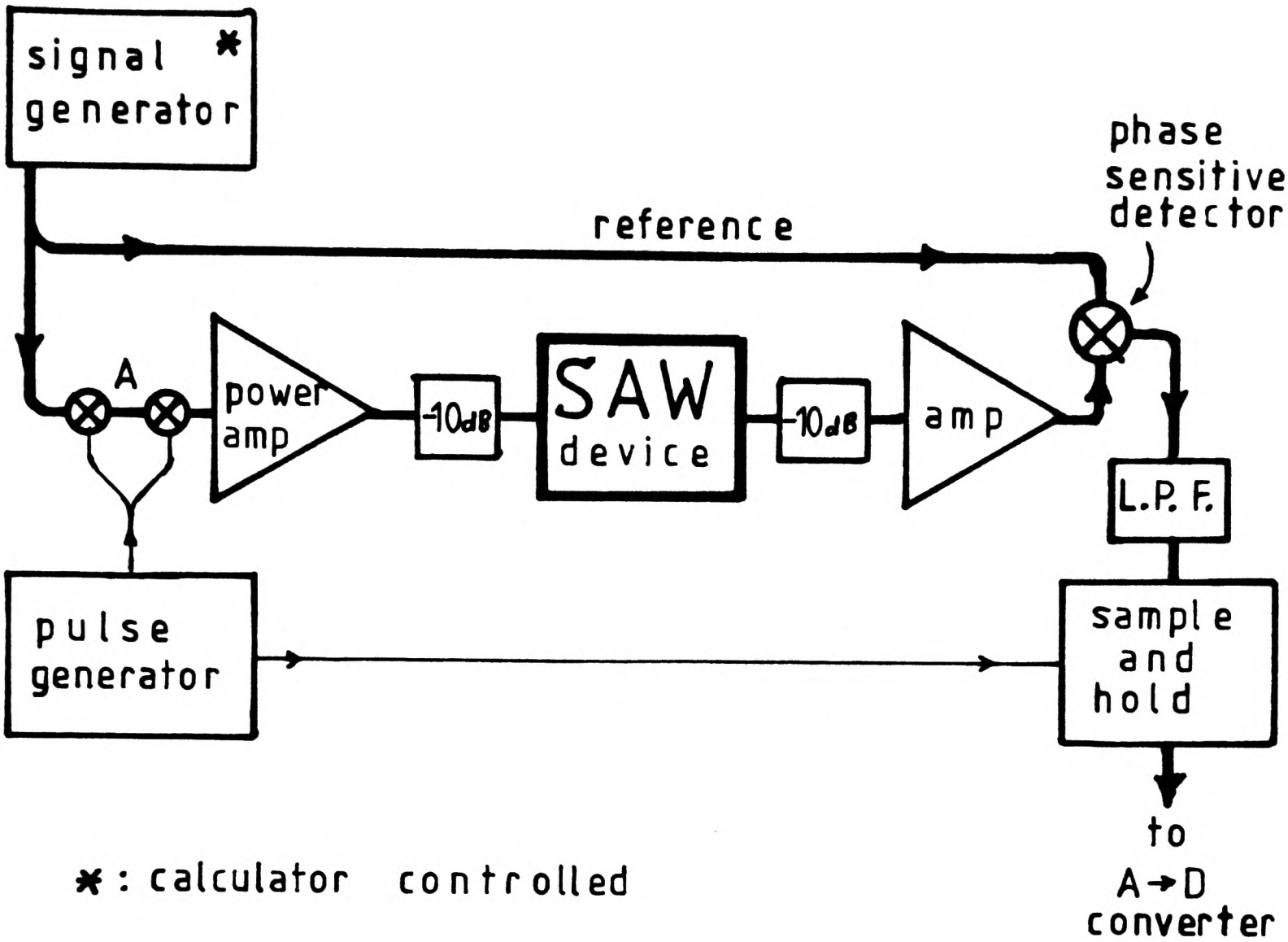


Fig. B1: Simplified block diagram of the phase measuring system for a RAC.

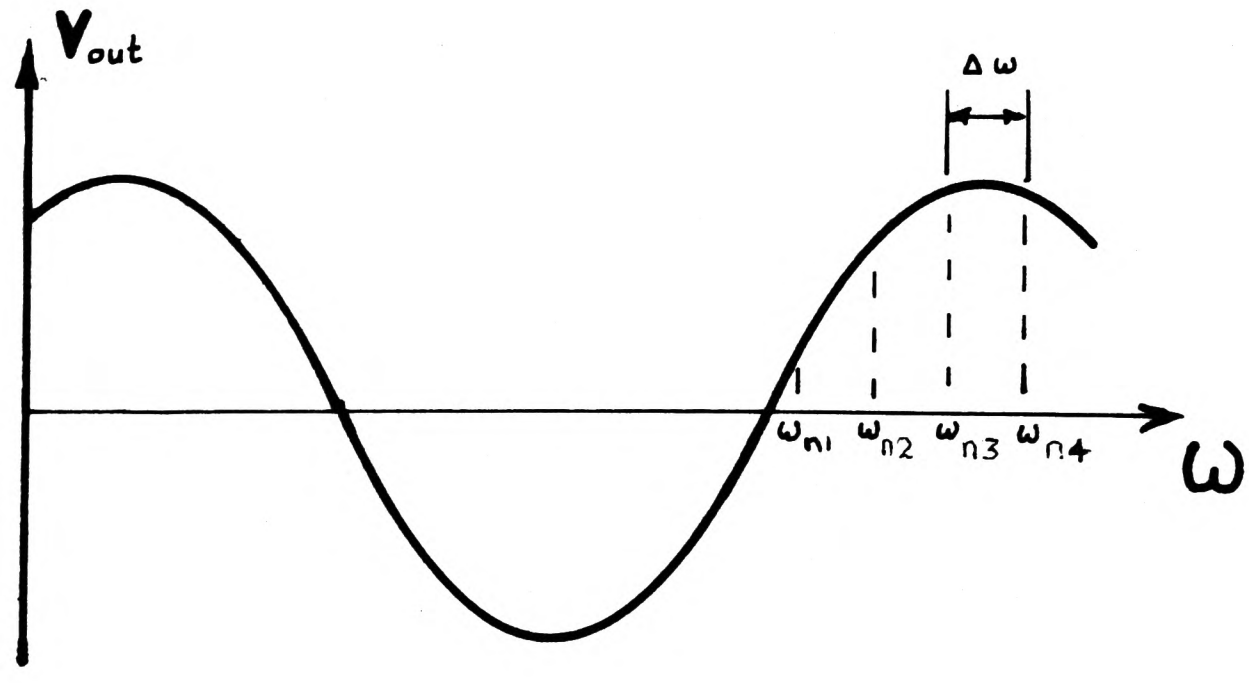


Fig. B2: Evaluation of a maximum point.

Appendix C

180° Reflections by Strips at Normal Incidence

This calculation deals with a charge varying in one dimension. The first stage, finding charge density given the incident wave, differs from its two-dimensional extension (chapter 3) as will be described. The second stage, finding the reflected wave from the charge, is given by equation (3.9), or ref.[30].

As in chapter 3, the incident wave is given by the real part of

$$V_o \exp(-jk_o x) \qquad \text{(C-1)}$$

where time variation of exp(jωt) is suppressed. Under a strip, the potential is V_d; the change in potential

$$\Delta V = V_d - V_o \exp(-jk_o x) \qquad \text{(C-2)}$$

is caused by charge redistribution on the strip.

Let the charge density on a strip be approximated by charges of q_j (coulombs/metre) distributed uniformly along infinitely long "cells" of width X (fig. C1). Using

$$\Delta V = \left(- \frac{1}{\pi(\varepsilon_o+\varepsilon)} \ln x + C \right) Q \qquad \text{(C-3)}$$

for a line charge Q, it can be found from integration that the potential change ΔV at point i=1 due to charge is given by

$$\Delta V_1 = \frac{1}{\pi_1(\varepsilon_0+\varepsilon)}\left[\left(-\ln\frac{1}{2e\eta}\right)q_1 + \text{------}\right.$$

$$\left. + \left(-\ln\left|\frac{x_1-x_j}{X\eta}\right| + \frac{1}{24}\left[\frac{X}{x_1-x_j}\right]^2\right) q_j + \text{----}\right] \qquad \text{(C-4)}$$

where **e**=2.718 and the constant of integration **C** has been replaced by $\ln(\eta)$. The coefficients

$$-\ln\left|\frac{x_1-x_j}{X\eta}\right| + \frac{1}{24}\left(\frac{X}{x_1-x_j}\right)^2$$

are found by power series expansion. In an isolated strip,

$$\sum_{j=1}^{n} q_j = 0 \qquad \text{(C-5)}$$

so that

$$\sum_{j=1}^{n}\left(-\ln\frac{1}{\eta}\right)q_j = 0 \qquad \text{(C-6)}$$

making η in (C-4) arbitary. (C-5), (C-4) and similar expressions for $\Delta V_2, \ldots \Delta V_n$, form (n+1) simultaneous equations from which the potential of the strip V_d and **n** values of q_j can be found as in chapter 3.

If the strip is not isolated, but in a periodic array of electrically unconnected strips, interactions can be dealt with without increasing the number of unknowns as described in chapter 3. If the strips are shorted, then (C-5) does not apply for an individual strip but V_d is zero, leaving **n** equations and **n** unknowns.

A small difficulty arises with shorted strips when the wavelength is not an integer number of periods. The potential change ΔV_i on a strip (fig. C2) is

$$\Delta V_i = \sum_{j=1}^{Nn} \left(\frac{-1}{\pi(\varepsilon_0+\varepsilon)} \ln\left|\frac{x_i - x_j}{X\eta}\right| \right) q_j \qquad (C-7)$$

where the summation is over **N** neighbouring strips, with a total of **Nn** cells. (The higher order terms $(X/|x_i-x_j|)^2/24$ are neglected in this discussion for simplicity, but included in the program.) The number of unknowns q_j is still **n**, since the charges at corresponding points on neighbouring strips are equal except for a known phase shift, and equations for **n** values of ΔV_i, all on the same strip, are used. It is necessary to take **N** strips such that

$$\sum_{j=1}^{Nn} q_j = 0 \qquad (C-8)$$

Otherwise, the constant of integration η must be infinite to make potential zero at infinite distance from the strips.

If a wavelength is not an integer number of periods, a region 1 (fig. C2) which has overall charge neutrality is difficult to find and may not contain an integer number of strips. An arbitary region 2 is taken (approximately 40 wavelengths wide in the program) which does have an integer number of strips. By choosing a value of η such that $X\eta=R$ in (C-7), the effect for the neglected charges in region 3 is minimised, since their coefficients $-\ln(|x_i-x_j|/X\eta)$ in the equations are small.

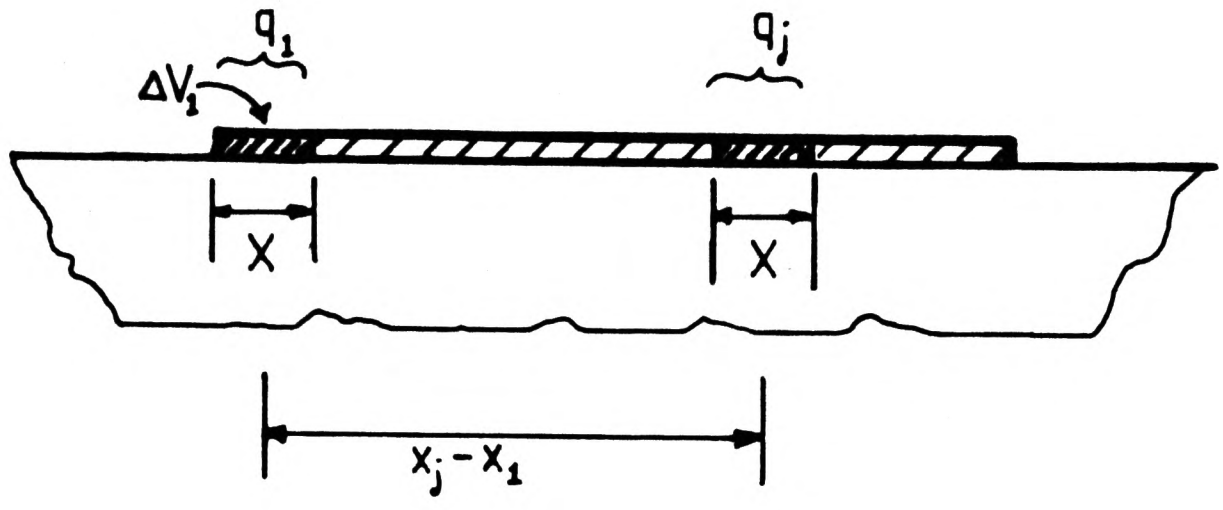


Fig. C1: Charges on an isolated strip.

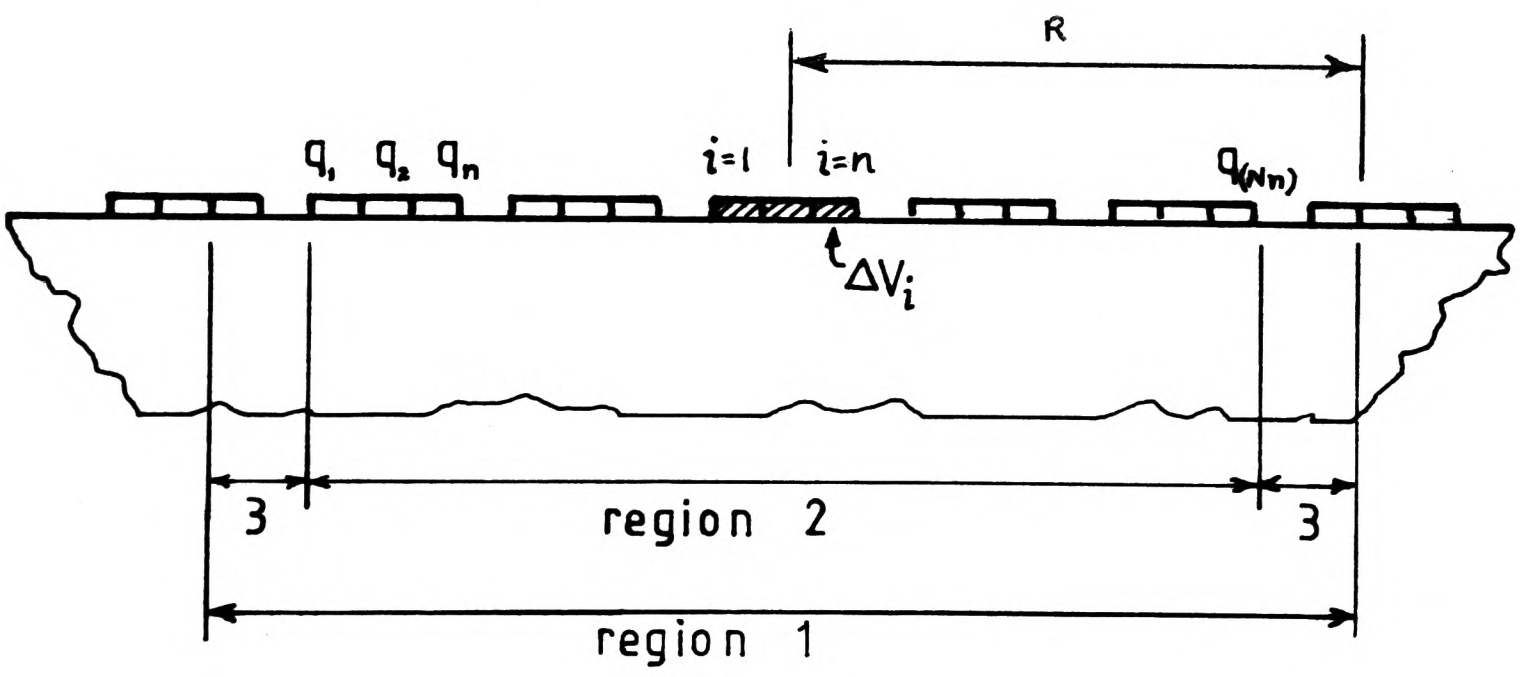


Fig. C2: Charges on a strip with near neighbours.

Appendix D

90° Reflections from Strips AT 45° Incidence

Calculations for reflections by strips (or very long dots) are identical to the computations for dots except for two simplifications, both in the first stage (finding charge density). Firstly, the strip is always at ground potential. The second difference is that charges on the 'cells' along a line parallel to the edge of the strip (fig. D1) are equal except for a phase delay factor $\exp(-jk_o z)$. So

$$\Delta V_i = -V_o \exp(-jk_o z_i) = \frac{1}{2\pi(\varepsilon_o+\varepsilon)} \sum_j \sum_m \frac{1}{l_{ijm}} q_{mj} \qquad \text{(D-1)}$$

$$-V_o \exp(-jk_o z_i) = \frac{1}{2\pi(\varepsilon_o+\varepsilon)} \sum_j \left\{ \sum_m \frac{1}{l_{ijm}} \exp\left[-jk_o(z_{mj}-z_j)\right] \right\} q_j \qquad \text{(D-2)}$$

In the program, **m** is summed over a region of ±5 wavelengths about $z=z_j$. Since the function

$$\frac{1}{l_{ijm}} \exp\left[-jk_o(z_{mj}-z_j)\right]$$

is hermitian about $z_{mj}=z_j$, the imaginary parts cancel in the summation and the function can be replaced by

$$\frac{1}{l_{ijm}} \cos\left[k_o(z_{mj}-z_j)\right]$$

The number of unknowns is greatly reduced since q_j varies only along one dimension. The summation in **m** as a function of **(i-j)** is performed once only, and the results stored in a look-up table, which is then re-used repeatedly to find the coefficients of q_j in the matrix equation.

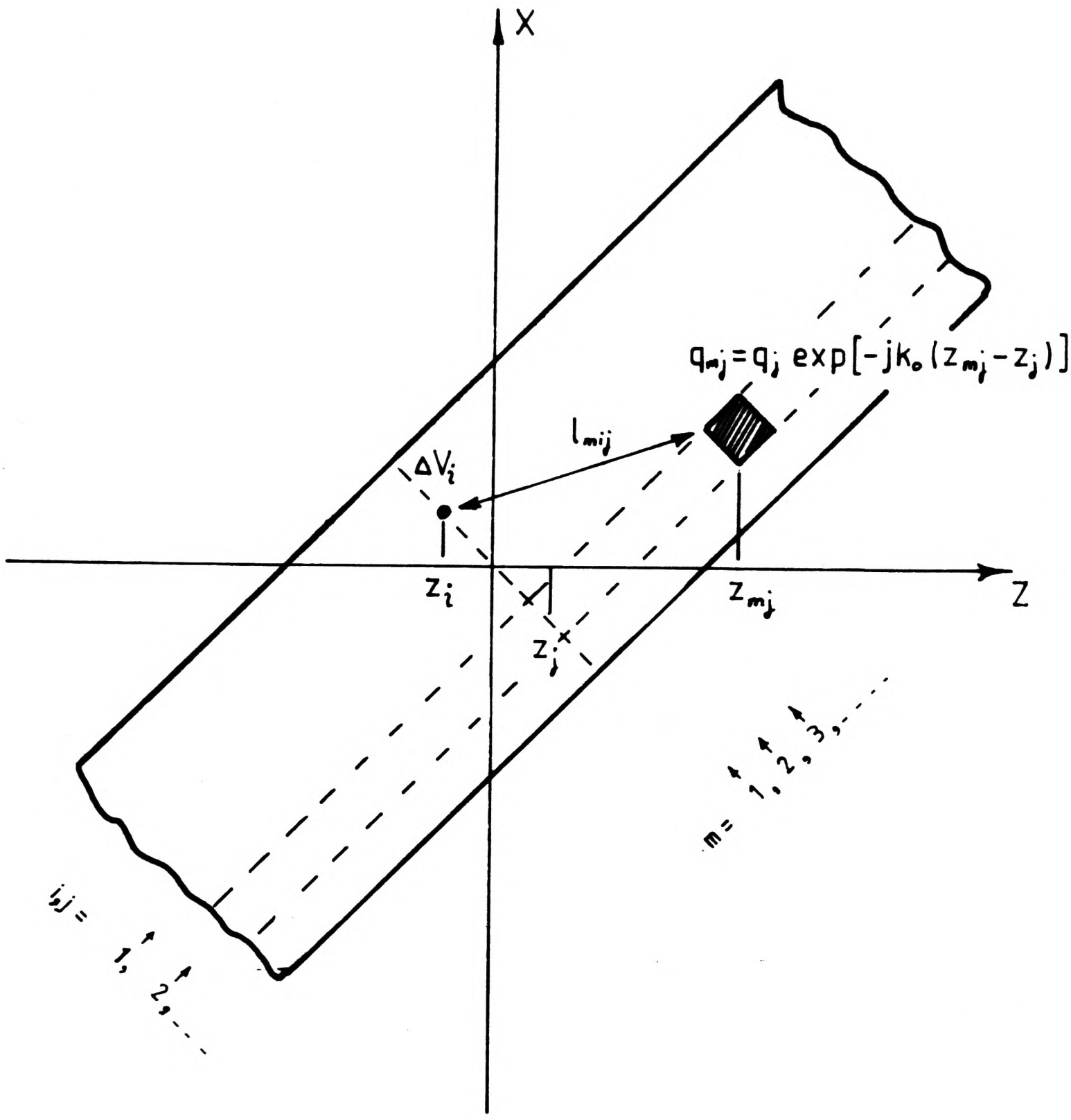


Fig. D1: Charges on a strip inclined at 45 degrees.

Appendix E

Frequency Respose of a Periodic Array in L-Path

An isotropic material is considered first. If small, identical scatterers at A,B and C have a spacing λ_o, where λ_o is slightly larger than λ, the wavelength of the incident wave (fig. E1), they scatter slightly out of phase, so the reflected wave **abc** is such that

$$\vec{Aa} + 2\lambda = \vec{AB}\vec{b} + \lambda = \vec{ABC} \qquad \text{(E-1)}$$

where $\vec{AB}\vec{b}$ represents the distance A to B to **b** in fig. E1. So

$$\alpha = \sin^{-1} \frac{\vec{Bb}}{\lambda_o} \qquad \text{(E-2)}$$

Since α is small, the following approximations are made:

$$\alpha = \tan^{-1} \frac{\vec{Bb}}{\lambda_o} \qquad \text{(E-3)}$$

and

Wavelength of reflected wave = $\lambda \cos(\alpha)$ (E-4)

The error in each case is the difference between unity and $\cos(\alpha)$. The approximations are illustrated in fig. E2, in which the slight change in direction of power flow is also neglected.

On an anisotropic material, it is assumed that

$$\frac{V_z}{\lambda_z} = f = \frac{V_x}{\lambda_x} \qquad \text{(E-5)}$$

although the reflected wave is not propagating exactly in the

X direction. This approximation also simplifies the problem of anisotropy; only the velocities in the **Z** and **X** directions need to be known.

Using the above approximation, the output of the L and U-path devices shown in fig. E3 is just the sum of contributions from all possible right-angle paths formed by the reflecting dots in the array. Performing that summation, and substituting

$$f_x = \frac{V_x}{l_x} \quad , \qquad f_z = \frac{V_z}{l_x} \tag{E-6}$$

(where f_x and f_z are parameters of the array, not the wave)

$$\sqrt{\left(\frac{P_{out}}{P_{in}}\right)_L} = r\sqrt{MN}\,\frac{\sin(N\pi f/f_z)}{N\sin(\pi f/f_z)}\,\frac{\sin(M\pi f/f_x)}{M\sin(\pi f/f_x)} \tag{E-7}$$

$$\sqrt{\left(\frac{P_{out}}{P_{in}}\right)_U} = r^2 M N \quad \frac{\sin(2N\pi f/f_z)}{2N\sin(\pi f/f_z)}\left(\frac{\sin(M\pi f/f_x)}{M\sin(\pi f/f_x)}\right)^2 \tag{E-8}$$

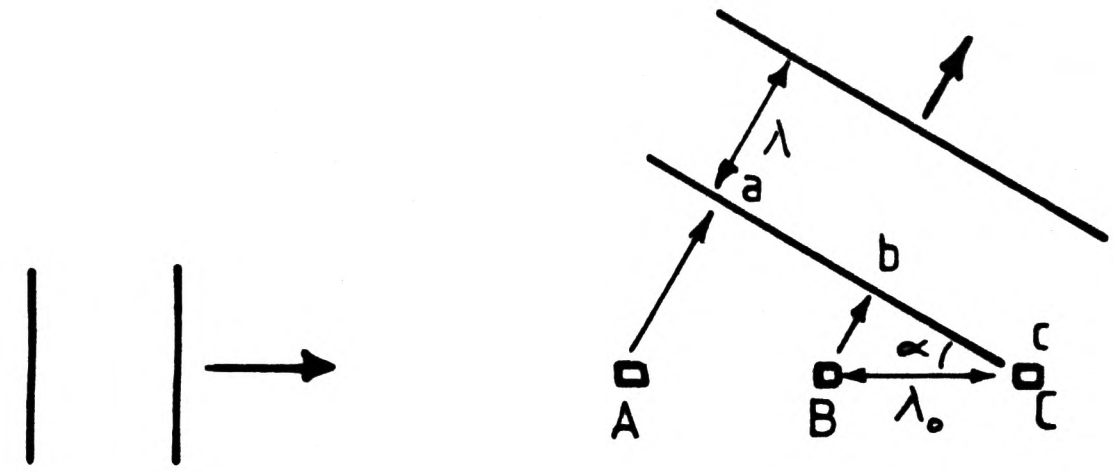


Fig. E1: Reflection from a single line of dots.

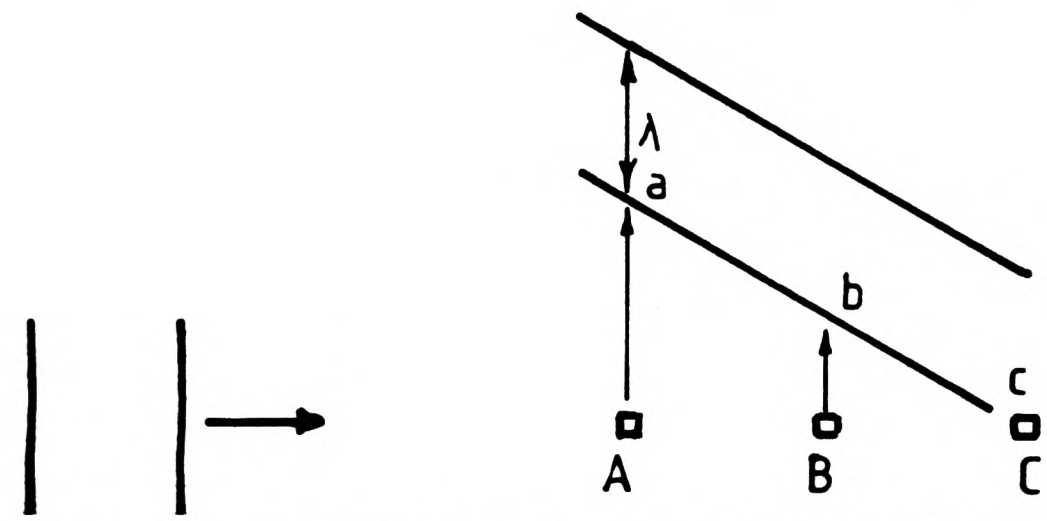


Fig. E2: Approximations for the configuration of fig. E1.

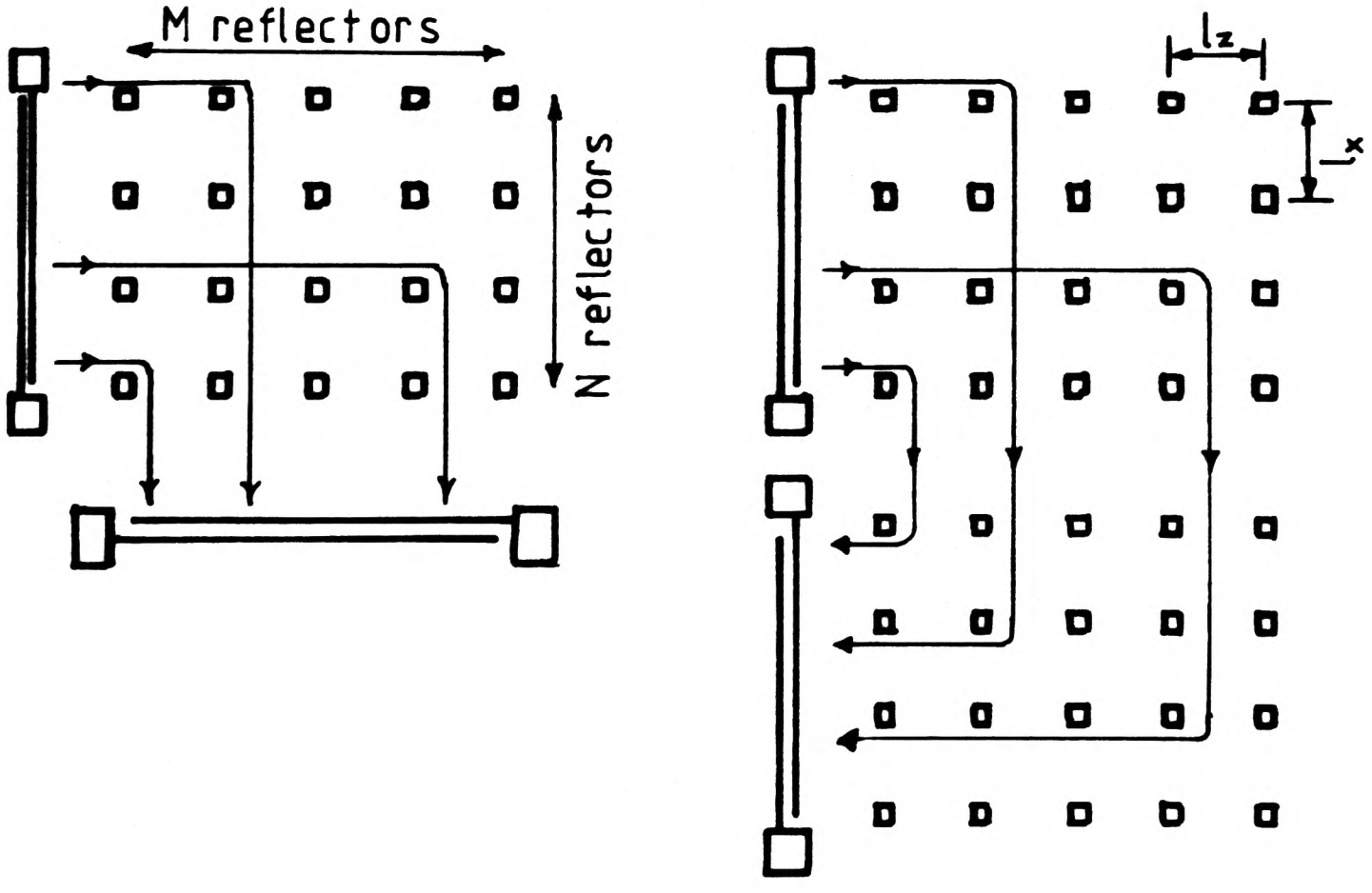


Fig. E3: "L" and "U"-path reflections in arrays.

Appendix F

Attenuation of an Incident Wave through Reflections

When a SAW is incident on a single reflector, some of its energy is diverted, so that the amplitude of the transmitted wave is smaller. In an array of reflectors, it is possible for the multiple reflections to be 180° out of phase with the transmitted beam; this also reduces the amplitude. In this appendix it is shown that the former effect is insignificant compared with the latter if the array is made up of a large number of weak reflectors.

It is assumed that reflectivity jr and the number of reflectors **N** is such that:

$$\begin{aligned} r &\ll 1 \\ Nr &\ll 1 \\ N &\gg 1 \end{aligned} \qquad \text{(F-1)}$$

Terms in r^3 and above can therefore be neglected. This implies that only double reflections need to be considered. For simplicity, only symmetric reflectors are dealt with; the reflection coefficient is then always imaginary [31].

Fig. F1 depicts an incident beam with amplitude **A** passing through **N** reflectors whose reflected beams do not overlap. The transmission coefficient through a single reflector is

$$\rho = \sqrt{1 - r^2} \qquad \text{(F-2)}$$

The amplitude of the transmitted wave is

$$T_1 = A\rho^N \simeq A\left(1 - \tfrac{1}{2}Nr^2\right) \qquad \text{(F-3)}$$

In fig. F2, the reflectors have been brought together so that they are spaced by an integer number of wavelengths. The reflected power is now $(ANr)^2$, so the amplitude of the transmitted wave is

$$T_2 \simeq A\left(1 - \tfrac{1}{2}N^2r^2\right) \qquad \text{(F-4)}$$

Fig. F3 shows that the additional amplitude loss can be associated with double reflections. (In fig. F3, phase delays due to path differences are neglected since path differences are an integer number of wavelengths).

On comparing (F-3) and (F-4), it becomes clear that the amplitude loss $(1-\rho)$ "per reflector" is relatively small. A calculation of amplitude loss based on ρ^N would be inappropriate, since the power in the reflected beam is more closely related to the level of double reflections.

A second simple case can be considered. If the strips in fig. F2 are positioned such that their reflected beams are out of phase and cancel, the beam is transmitted without attenuation. A calculation of amplitude loss based on ρ^N is again inappropriate. Numerical calculations for specific examples have been made (using equation 2.4). In these cases, the term due to double reflection is negative and cancels with the loss associated with ρ^N.

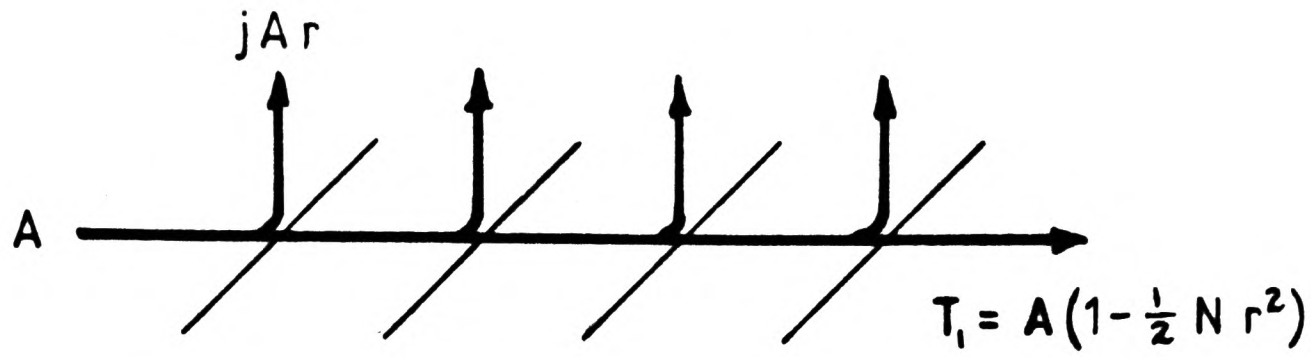


Fig. F1: N Reflectors whose reflected beams do not overlap.

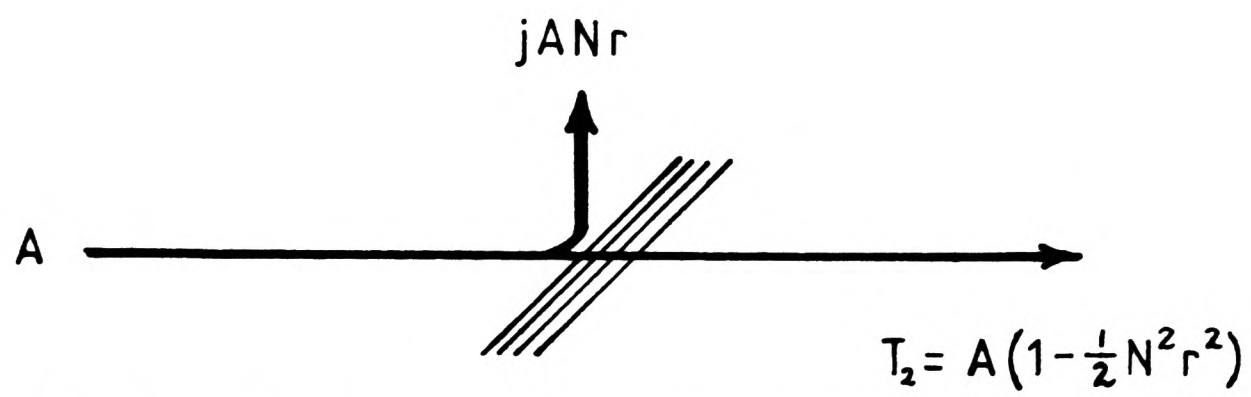


Fig. F2: N Reflectors whose reflected beams completely overlap.

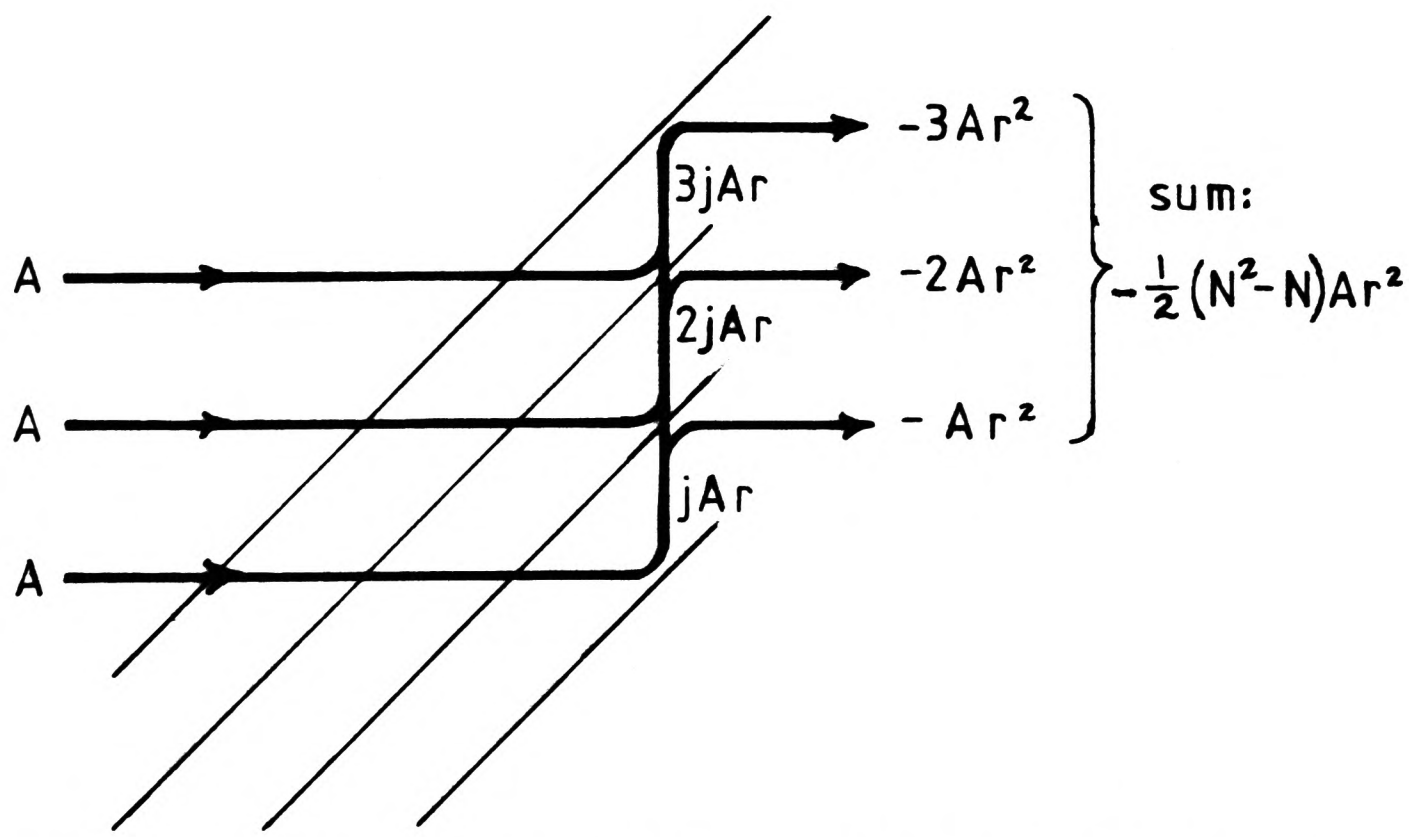


Fig. F3: Double reflection paths for the N reflectors in fig. F2. (N=4)

Appendix G

Power Lost By Reflections in a Chirped Array

It is shown here that the fraction of incident power reflected by a chirped array is $(rN_e)^2$, where jr is the reflectivity per reflector (strip, groove or row of dots), and N_e the "effective" number of reflectors as defined by (5.25a). This value is independent of the aperture of the array.

This is obvious when the aperture is very large compared with the width of the synchronous region (fig. G1(i)). For an incident amplitude **A**, the reflected wave has amplitude

$$R_1 = A N_e r \qquad \text{(Neglecting terms in } r^3 \text{ and above)} \qquad \text{(G-1)}$$

inside a region of **z** with width equal to D_1, and the power removed from the incident beam is as given above.

When **D** is small (fig. G1(ii)), the magnitude of the reflected wave at **q** is smaller than $(AN_e r)$, because there are fewer reflectors intersecting the vertical line through **q**. However, at **p**, where the reflectors do not reflect exactly synchronously, the reflected wave is larger than for the wide-aperture array, because the narrow array is less frequency-selective. By numerical integration, the fraction of power reflected

$$\frac{P_{reflect}}{P_{in}} = \frac{1}{A^2 D} \int |R_2|^2 \, dz \qquad \text{(G-2)}$$

can be shown to be independent of aperture, and equal to $(N_e r)^2$. This will be outlined later. It should be noted that

the reflected wave along **pqs** is not a straight-crested wave; after being reflected by the second array (not shown), the components do not add exactly in phase at the output transducer, so the response of the whole device is indeed dependent on aperture, as given by (5.25), unlike the power loss being discussed here.

It will next be shown that, far away from the synchronous region, the amplitude of the transmitted wave is uniform. The lower section of the array in fig. G1(i), of width $x=D_2$, is identical to the narrow array (Fig. G2(ii)), so that power transmitted by the lower section of the wide array is

$$\int_0^{D_2} \left| T_1(x) \right|^2 dx = A^2(1 - N_e^2 r^2) D_2 \qquad \text{(G-3)}$$

or, differentiating,

$$\left| T_1(x) \right|^2 = A^2 (1 - N_e^2 r^2) \qquad \text{(G-4)}$$

So the amplitude of the transmitted wave is uniform.

Finally, it is indicated that the phase is also uniform and not affected by double reflections. As illustrated in Appendix F, power lost by the array is closely linked to the level of double reflections. A down-chirp array is assumed. For any double reflection path at **p** (fig. G2) for which the increase in path length is slightly less than an integer number of wavelengths, there is a double reflection path at **s**, in which the path length is slightly more than an integer number of wavelengths, so the out-of-phase conponents cancel.

Outline of Calculations to Find Power Reflected

From (5.22) and (5.25a),

$$Z_n = Z_o + \frac{v f_1 T}{2B}\left[1 - \sqrt{1 - \frac{4B}{f_1^2 T} n}\right] \qquad \text{(G-5)}$$

$$N_e = f\sqrt{\frac{T}{2B}} \qquad \text{(G-6)}$$

Since $\lambda_i = \frac{dz}{dn}$ and $f = v/\lambda$, the value of n_s corresponding to the center of the synchronous region for frequency f is found from (G-5) and (G-6) to be

$$n_s = \frac{1}{2} N_e^2 \left[\left(\frac{f_1}{f}\right)^2 - 1\right] \qquad \text{(G-7)}$$

An index transformation will be made so that reflectors are numbered m and have positions l_m, both measured from the centre of the synchronous region (fig. G3). The transformations are

$$n = n_s + m \qquad \text{(G-8)}$$

$$Z_n = Z_s + l_m \qquad \text{(G-9)}$$

By manipulating (G-5) to (G-9), it can be shown that

$$l_m = \lambda N_e^2 \left[1 - \sqrt{1 - \frac{2}{N_e^2} m}\right] \qquad \text{(G-10)}$$

$$l_m \simeq \lambda\left[m + \frac{1}{2}\left(\frac{m}{N_e}\right)^2\right] \qquad m << N_e^2 \qquad \text{(G-11)}$$

(G-11) makes use of a binomial expansion of the square root.

Neglecting a constant phase term, the reflected wave is given by

$$R = \sum_{m_1}^{m_2} A\, r_m \exp[j 2\pi l_m/\lambda] \qquad \text{(G-12)}$$

Reflectivity per reflector jr_m is assumed constant. The limits of the summation are indicated in fig. G3. (G-11) is subtituted into (G-12), also using $\exp(j2\pi m)=1$ for integer values of **m**.

$$R = A\, r \sum_{m_1}^{m_2} \exp\left(j\pi\left(\frac{m}{N_e}\right)^2\right) \qquad \text{(G-13)}$$

$$R \simeq A\, r \int_{(l-D/2)/\lambda}^{(l+D/2)/\lambda} \exp\left(j\pi\left(\frac{m}{N_e}\right)^2\right)\, dm \qquad \text{(G-14)}$$

Substituting $\eta=(\sqrt{2})\,m/N_e$ reduces (G-14) to a standard form:

$$R \simeq A\,\frac{N_e r}{\sqrt{2}} \int \exp\left(j\frac{\pi}{2}\eta^2\right)\, d\eta \qquad \text{(G-15)}$$

$$\simeq A\,\frac{N_e r}{\sqrt{2}}\left[\, E\left(\sqrt{2}\,\frac{l}{\lambda N_e} + \frac{1}{\sqrt{2}}\,\frac{D}{\lambda N_e}\right) - E\left(\sqrt{2}\,\frac{l}{\lambda N_e} - \frac{1}{\sqrt{2}}\,\frac{D}{\lambda N_e}\right)\,\right] \qquad \text{(G-16)}$$

where **E( )=C( )+jS( )**, and **C( )** and **S( )** are Fresnel integrals.

The final stage, evaluating the fraction of power lost

$$\frac{1}{A^2 D}\int |R|^2\, dl = \frac{1}{A^2\left(\frac{D}{\lambda N_e}\right)} \int \left| R\left(\frac{l}{\lambda N_e}, \frac{D}{\lambda N_e}\right)\right|^2 d\left(\frac{l}{\lambda N_e}\right) \qquad \text{(G-17)}$$

is done numerically. The limits of the integral can be assumed to be any arbitary large value because **R** is both small

and rapidly oscillating far away from the synchronous region. After integration, the dummy variable $(1/\lambda N_e)$ disappears and power reflected is a function of $(D/\lambda N_e)$ only, but is found to be independent of this latter variable, being equal to $(rN_e)^2$.

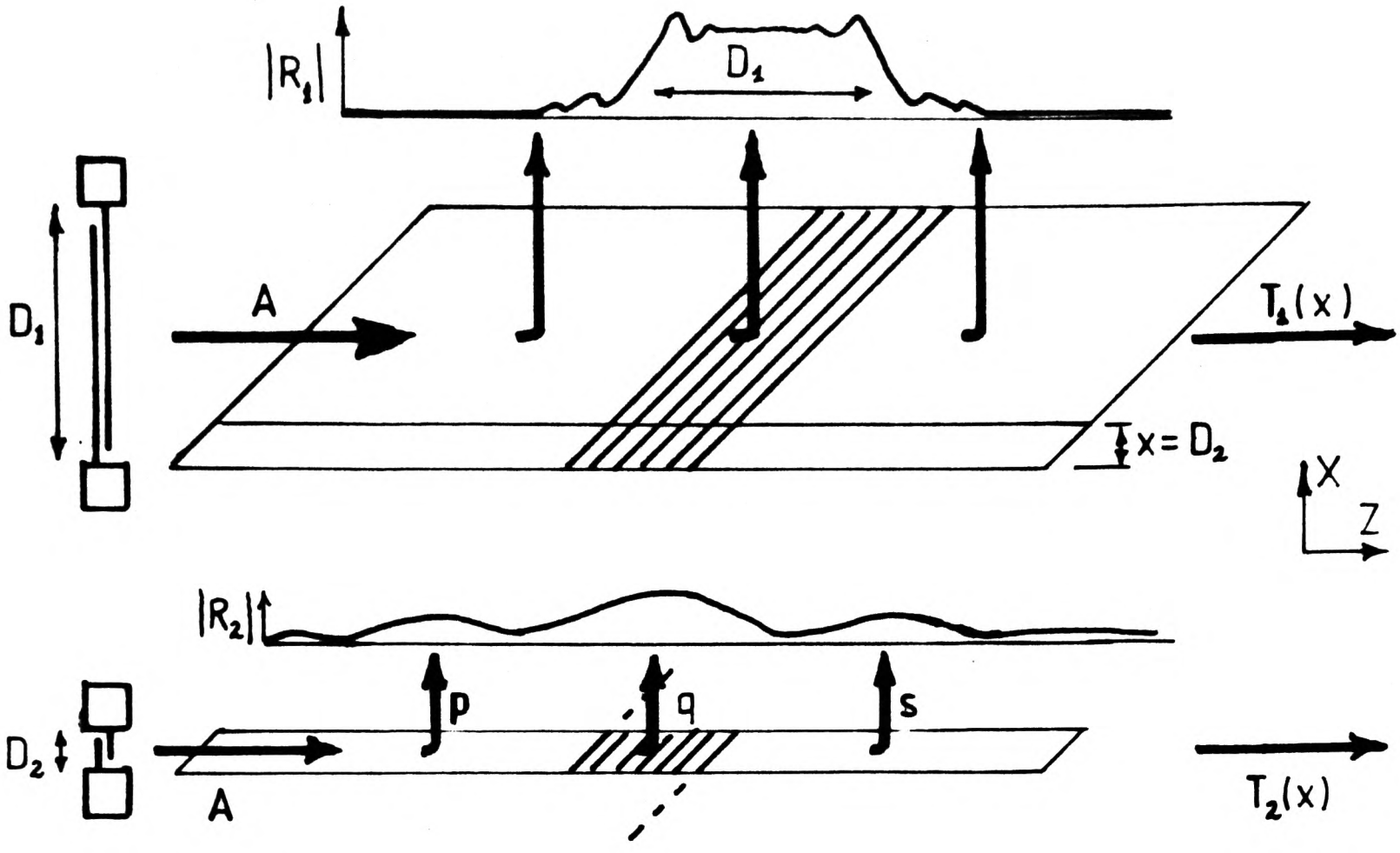


Fig. G1: Reflection in an array of a RAC with (i) wide-aperture and (ii) narrow-aperture.

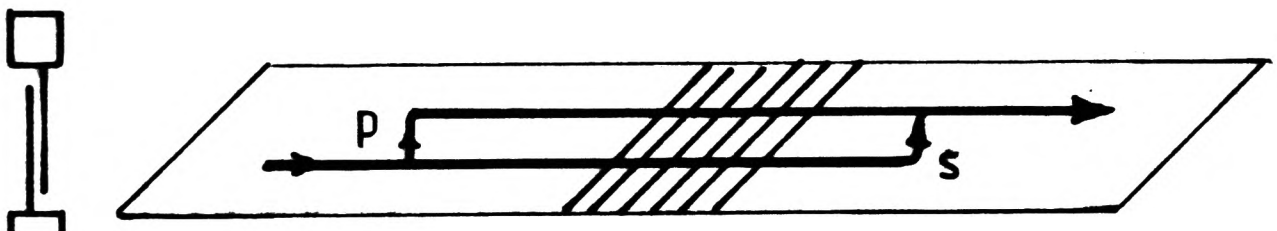


Fig. G2: Double reflection paths in a RAC array.

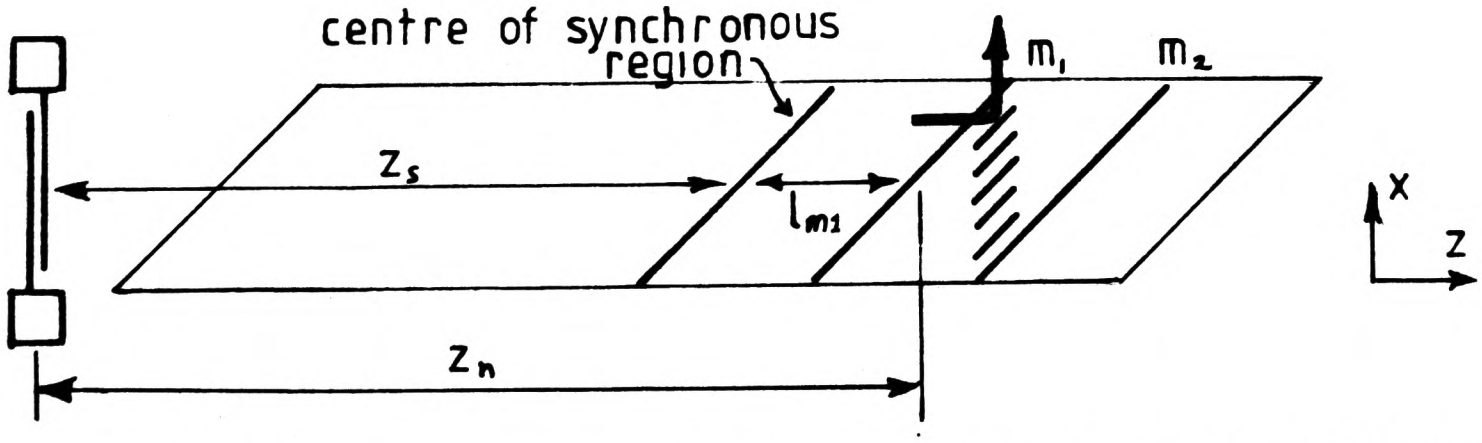


Fig. G3: Transformation of indices used in the calculation of the magnitude of the reflected wave.

Appendix H

The Effect of Position Errors in a RAC on its Response

The position errors of all the rows in a RAC, which usually vary slowly from row to row, can be expressed as a function **Z(z)**, where **z** is the position of a row and **Z(z)** is the position error. The instantaneous frequency associated with any row is clearly affected by **Z'(z)** (i.e. $\frac{d}{dz}Z(z)$), and hence the frequency response might be expected to be affected by both **Z(z)** and **Z'(z)**. It is shown here that the effect of **Z'(z)** is small.

A linear down-chirp RAC with the above error has an impulse response

$$f(t) = a(t) \exp j\left[\omega_o t - \tfrac{1}{2}\mu t^2 + e(t)\right] \qquad \text{(H-1)}$$

A relation between **e(t)** and **Z(z)** is required. Since the path is folded, the strip at position **z** affects the impulse reponse at time **t** such that **z=vt/2**. The impulse response is shifted by a time **Δt=(2/v)Z(vt/2)**. So an ideal response

$$f_o(t) = a(t) \exp j\left[\omega_o t - \tfrac{1}{2}\mu t^2\right] \qquad \text{(H-2)}$$

is modified to

$$f(t) = a\left(t - \tfrac{2}{v}Z\left(\tfrac{vt}{2}\right)\right) \exp j\left[\omega_o\left[t - \tfrac{2}{v}Z\left(\tfrac{vt}{2}\right)\right] - \tfrac{1}{2}\mu\left[t - \tfrac{2}{v}Z\left(\tfrac{vt}{2}\right)\right]^2\right]$$

$$f(t) = a\left(t - \frac{2}{v}Z\left(\frac{vt}{2}\right)\right) \exp j\left[\omega_0 t - \frac{1}{2}\mu t^2 - (\omega_0 - \mu t)\frac{2}{v}Z\left(\frac{vt}{2}\right) + \frac{2\mu}{v^2}Z^2\left(\frac{vt}{2}\right)\right] \quad \text{(H-3)}$$

Comparing (H-1) and (H-3), taking only linear terms,

$$e(t) \simeq -2\frac{\omega_i}{v}Z\left(\frac{vt}{2}\right) \quad \text{(H-4)}$$

where ω_i is the instantaneous frequency of the impulse response at any given time. The shift in the amplitude term is usually insignificant in practical devices.

The frequency response of the RAC $F(\omega)$ is found by using a method similar to finding the frequency spectrum of a signal (chapter 5, section 1, (5.7) and (5.7a)).

$$F(\omega) = \sqrt{2\pi}\frac{a(\tau + e(\tau)/\omega)}{\sqrt{\mu - e''(\tau)}} \exp j\left[\frac{(\omega-\omega_0)^2 - (e'(\tau))^2}{2\mu} - \frac{\pi}{4} + e(\tau)\right] \quad \text{(H-5)}$$

where, from (5.7a), τ is given implicitly by

$$\tau = \frac{-\omega + \omega_0 + e'(\tau)}{\mu} \quad \text{(H-5a)}$$

Provided that $a(t)$ is slowly varying and $e''(t)$ is small, the amplitude is not affected significantly by the error $e(t)$. The only first order term in the phase is $e(t)$; the gradient $e'(t)$ has only a second order effect. The phase error is therefore proportional to the position error of those strips which reflect synchronously at the frequency in question, and this fact is used when correcting the phase of a device. The

error gradient does not enter into the calulation. The group delay, given by the derivative of -arg(F(ω)) in (H-5),

$$t_g = -\frac{d\phi}{d\omega} = \frac{-\omega + \omega_0 + e'(\tau)}{\mu} \qquad \text{(H-6)}$$

is indeed proportional to the error gradient, but group delay is not used directly when making the above corrections.

Appendix I

Waves Reflected by a Distribution of Scatterers

The Amplitude of the Scattered Wave

The first part of this appendix derives a relationship between the Fourier transform of a distribution of scatterers and the amplitude of the reflected wave. Two dimensional space, appropriate for SAWs, is assumed. In fig. I1, **S** is a region of scatterers with density $h(\underline{l})$, where $\underline{l}$ is a position vector. **ds** is a small area. **p** is the position vector of a point far away from the region **S**. For simplicity, the space is taken to be isotropic.

It is assumed that an incident wave, equal to the real part of $A \exp j(\omega t - \underline{k}_{in} \cdot \underline{l})$ is scattered isotropically by scatterers in area **ds**. Since **P** is far away from **S**, the component of the scattered wave which arrives at **P** has wave vector $\underline{k}_{out}$ parallel to $(\underline{p}-\underline{l})$. The scattered wave at **P** due to scatterers in **ds** is proportional to

$$\exp j[\,\omega t - \underline{k}_{in} \cdot \underline{l}\,]\; h(\underline{l})\; \frac{1}{|\underline{p}-\underline{l}|}\; \exp j[\,-|\underline{k}_{out}||\underline{p}-\underline{l}|\,] \qquad \text{(I-1)}$$

Integrating this over the region **S**, and neglecting the $1/(\underline{p}-\underline{l})$ factor because it is slowly varying, the reflected wave is proportional to

$$\int_S \exp j[\omega t - \underline{k}_{in} \cdot \underline{l}]\; h(\underline{l})\; \exp j[-\underline{k}_{out} \cdot \underline{p} + \underline{k}_{out} \cdot \underline{l}]\; ds$$

$$= \exp j[\omega t - \underline{k}_{out} \cdot \underline{p}]\; H(\underline{k}_{in} - \underline{k}_{out}) \qquad \text{(I-2)}$$

where $H(\underline{k})$ is the spatial Fourier transform of $h(\underline{l})$. There is

an additional condition, which in isotropic space is that

$$\left|\underline{k}_{out}\right| = \left|\underline{k}_{in}\right|$$

If $\mathbf{H}(\underline{k})$ can be approximated as a set of delta functions, each in a different position $-\underline{k}_g$ in Fourier transform or "**k**"-space, each delta function can be associated with a "scattering vector" $\underline{k}_g$. (The minus sign is necessary above to conform with the notation of ref. [47].) Each scattering vector defines a point in **k**-space where $\mathbf{H}(-\underline{k})$ is non-zero which by (I-2) can give rise to a non-zero reflection. The conditions for a reflection are then that $\underline{k}_{out}$ exists such that

$$\underline{k}_{out} = \underline{k}_{in} + \underline{k}_g \qquad \text{(I-3)}$$

and

$$\left|\underline{k}_{out}\right| = \left|\underline{k}_{in}\right| \qquad \text{(I-4)}$$

Power Lost by a Randomised Array

An example of a "randomised" array is one in which the rows of reflectors are in predetermined postions, but the position of each scatterer along a row is random (or pseudo-random). The number of scatterers is unchanged by randomisation (only their postions are changed), so

$$\int_S \left| h(z,x)\right|^2 dz\, dx \qquad \text{(I-5)}$$

is constant. By a standard Fourier transform relationship, the quantity

$$\int_{\text{all k-space}} \left| H(k_z, k_x) \right|^2 \, dk_z \, dk_x \qquad \text{(I-6)}$$

is also invariant under randomisation.

Suppose that the input wave consists of a wide band of values of $\underline{\mathbf{k}}_{in}$, all in the same direction. Each scattering vector within the small area **A** in **k**-space (fig. I2) reflects its own frequency under the scattering conditions (I-3) and (I-4). The total power scattered corresponding to **A** is proportional to

$$\int_{A} \left| H(k_z, k_x) \right|^2 \, dk_z \, dk_x \qquad \text{(I-7)}$$

From (I-6), the average value of $|\mathbf{H}(\underline{k})|^2$ is not reduced by randomisation. Neither will randomisation move the scattering vectors (non-zero values of $\mathbf{H}(\underline{k})$) out of any particular region **A**, except by fortuitous accident. (The regions **A** of interest are those which contain scattering vectors $\underline{\mathbf{k}}_g$ which satisfy the scattering conditions at frequencies inside the pass band of a device.)

It can be seen that although randomisation will convert each scattering vector $\underline{\mathbf{k}}_g$ into either a large number of small ones or a continuous distribution of $\underline{\mathbf{k}}_g$'s (neither proved here), the total power scattered is in general not significantly reduced. It is however possible that randomisation using specially-chosen probability distribution functions may be useful. These could be studied making use of the fact that $|\mathbf{H}(\underline{k})|^2$ is the Fourier transform of the autocorrelation of $\mathbf{h}(\underline{l})$.

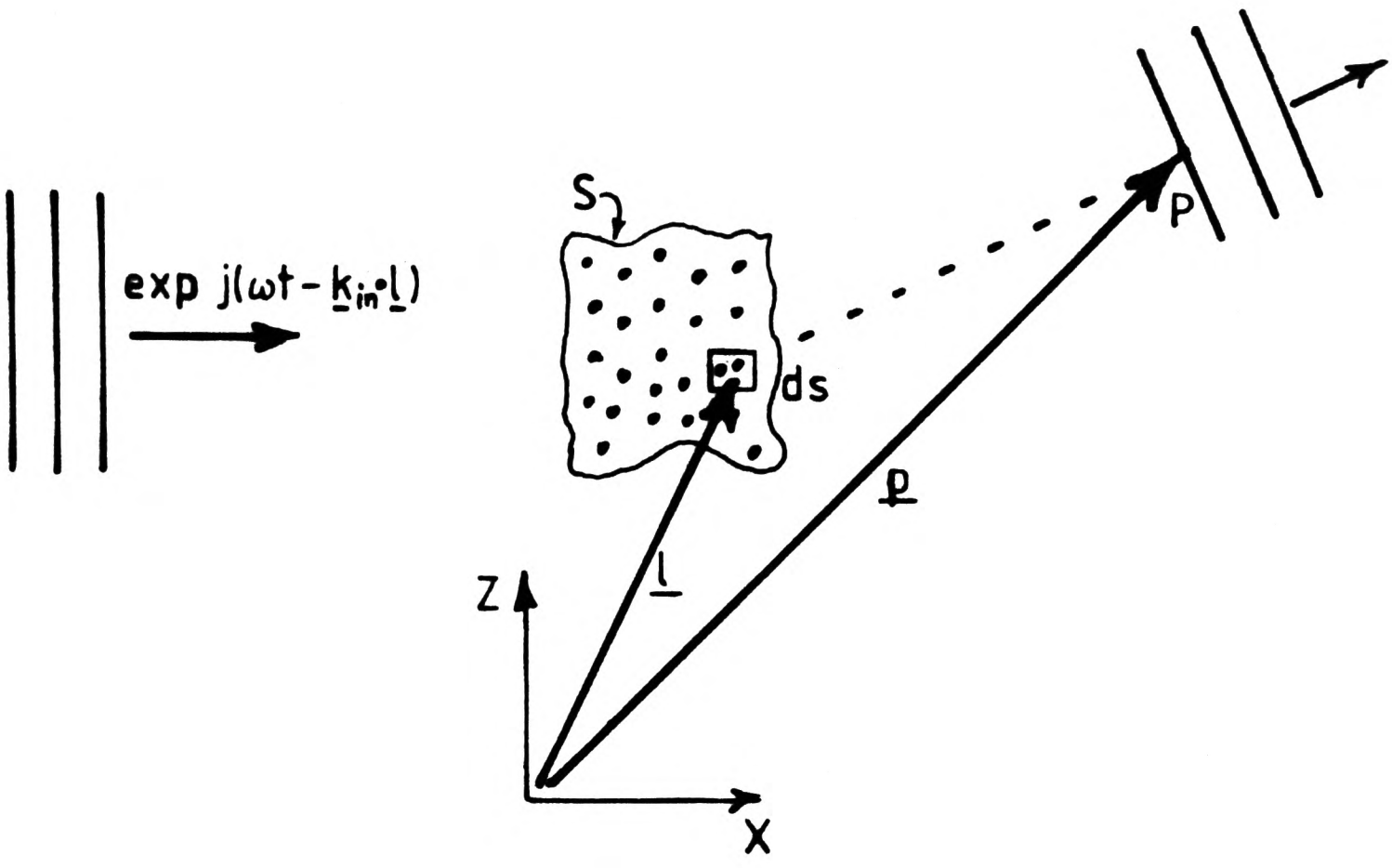


Fig. I1: Reflection from a distribution of scatterers.

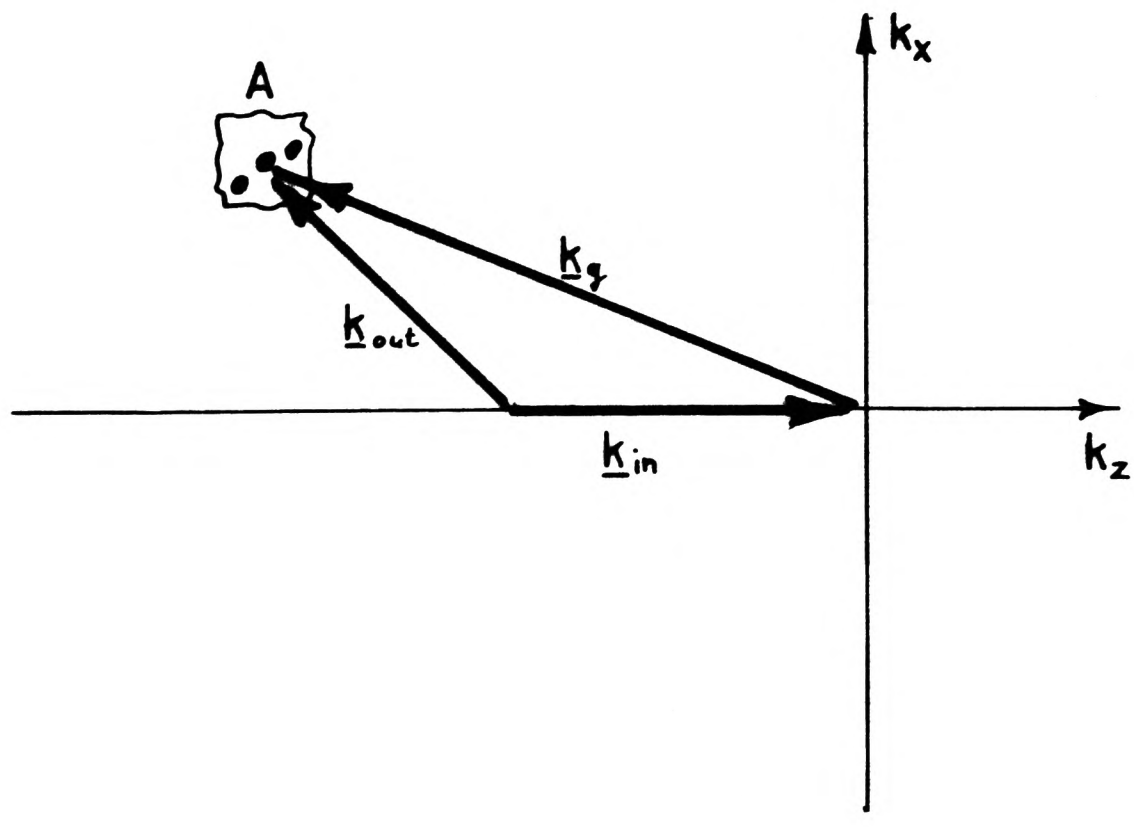


Fig. I2: Scattering from those scattering vectors which are inside a small region A.

References

1. Ameri S, Ash E, Htoo U, Murray D, Wickramasinghe H
   Private communication

2. Bertoni H L
   Right-angle reflection of SAW from an array of conducting strips
   IEEE trans. SU-30, no. 2, Mar 1983

3. Bloch P D, Paige E G S, Stove A G
   Selective reflection of SAWs by periodic dot arrays
   1979 IEEE ultrasonics symposium proceedings, pg 687

4. Bloch P D, Paige E G S, Solymar L
   Reflection of surface acoustic waves form arrays of strips
   1978 IEEE ultrasonics symposium proceedings

5. Boroson D M, Oates D E
   Experimental and theorectical analysis of temperature dependence of wideband SAW RAC devices on quartz
   1981 ultrasonics symposium proceedings

6. Bracewell R N
   The fourier transform and its applications
   McGraw-Hill publications

7. Chapman R E, Chapman R K, Morgan D P, Paige E G S
   Weighted in-line reflective array devices
   1979 IEEE ultrasonics symposium proceedings

8. Cook C E, Bernfeld M,
   Radar signals
   Academic Press

9. Datta S, Hunsinger B J
   An analytical theory for the scattering of SAWs by a single electrode in a periodic array on a piezoelectric substrate
   J Appl Phys (USA), Vol 51, No 9, Sep 80, pp 4817-23

10. Dunrowicz C, Sandy F, Parker T
    Reflection of SAWs from periodic discontinuities
    1976 IEEE ultrasonics symposium proceedings, pg 386 ff

11. Gerrard H M, Otto O W, Weglein R D
    Development of a broadband reflective array 10000:1 pulse compression filter
    1974 IEEE ultrasonics symposium proceedings, pg 197 ff

12. Godfrey J T, Grauling C H, Moore R A, Nothnick C E, Armatrout R J
    Improvements in metallised RAC design
    1977 IEEE ultrasonics symposium proceedings

13. Hartman C S, Kansy R J, Wright P V, Garber E M
Single phase unidirectional SAW transducers
1982 IEEE ultrasonics symposium proceedings

14. Hohkawa K, Yoshikawa S
Multichannel SAW filters employing two-dimensionally periodic RDA
Transactions of the IECE of Japan, Vol E64, No 12, Dec 81

15. Huang F, Paige E G S
Influence of size and shape on reflection of SAWs by thin aluminium dots on lithium niobate
Electronics letters, Vol 18, No 5, Mar 82

16. Huang F, Paige E G S
Reflection of SAWs by thin metal dots
1982 ultrasonics symposium proceedings

17. Kitano T, Nishikawa K
SAW right-angle reflection coefficient from metal dot array
Proc. of the 1st symposium on ultrasonic electronics, Tokyo, 1980

18. Kitano T, Nishikawa K, Asakawa K
Amplitude weighted metal strip reflective array compressor
1977 IEEE Ultrasonics Symposium Proceedings, pg 585 ff

19. Kitano T, Nishikawa K
SAW right-angle reflection coefficient from metal strip array reflector on Y-cut LiNbO3
Electronics Letters, Vol 13, No 5, pp 136-7, Mar 77

20. Yamanouchi K, Shibayama K
Propagation and amplification of Rayleigh waves and piezoelectric leaky surface waves in LiNbO3
Journal of Applied Physics, Vol 43, No 3, Mar 72

21. Marshall F G, Paige E G S
Mode conversion in SAW reflective arrays
Electronics letters, vol 10, no 8, 18 Apr 74

22. Marshall F G, Paige E G S, Young A S
Amplitude weighting of SAW reflecting array structures
1974 IEEE ultrasonics symposium proceedings, pg 202-4

23. Martin T A
The IMCON pulse compression filter and its applications
IEEE transactions, MTT-21, No 4

24. Matthaei G L, Barman F, Savage E B
SAW reflecting arrays
Electronics Letters, Vol 12, No 21, Oct 76

25. Matthaei G L, Barman F
SAW resonators using low-loss "waffle-iron" reflectors
1976 IEEE ultrasonics symposium proceedings, pp 415 ff

26. McCue J J G
A note on the Hamming weighting of linear FM pulses
Proceedings of the IEEE, Vol 67, No 11, Nov 79

27. Milsom R F
Analysis of generation and detection of surface and bulk acoustic waves by interdigital transducers
IEEE transactions SU-24, no 3, May 77

28. Milsom R F, Redwood M, Reilly N H C
The interdigital transducer
Surface wave filters, Matthews (Ed), Wiley-interscience

29. Milsom R F,
Private communication on several occasions

30. Morgan D P
Quasi-static analysis of generalised SAW transducers using the Green's function method
IEEE transactions, SU-27, no 3, May 80

31. Otto O W
Multiple reflections in SAW reflective arrays
IEEE transactions SU-22, No 4, Jul 1975

32. Paige E G S, Stove A G, Woods R C
SAW reflection from aluminium strips on LiNbO3
1981 IEEE Ultrasonics symposium proceedings

33. Peach R C
Private communication

34. Penunuri D, Thoss J L
Low loss, low spurious RDA filter using three-phase transducers
1981 IEEE Ultrasonics symposium proceedings, pg 23

35. Riha G, Stoker H R, Veith R, Bulst W E
RAC filters with position weighted metallic strip arrays
1982 ultrasonics symposium proceedings

36. Solie L
Reflective dot array devices
1977 IEEE Ultrasonics symposium proceedings, pg 579

37. Solie L
A SAW filter using a reflective dot array
1976 IEEE Ultrasonics symposium proceedings, pg 309-12

38. Solie L
Surface acoustic wave reflective dot array
Applied physics letters, vol 28, no 8, Apr 1976

39. Solie L
The development of high-performance RDA devices
1979 IEEE Ultrasonics symposium proceedings, pg 682-88

40. Stove A G
D Phil thesis
Oxford

41. Sun B S
Private communication

42. Szabo T L, Slobodnik A J
The effect of diffraction on the design of SAW devices
IEEE transactions, SU-20, pp 240-51, 1973

43. Thoss J L, Penunuri D, Thostenson M
Implementation of reflective array matched filters for radar applications
1981 IEEE ultrasonics symposium proceedings, pg 63-8

44. Van de Vart H, Solie L
A SAW pulse compression filter using the RDA
Applied Physics Letters, vol 31, pg 1, Jul 1977

45. Williamson R C
Propertices and applications of reflective-array devices
Proceedings of the IEEE, vol 64, no 5, May 1976

46. Williamson R C, Smith H I
The use of SAW reflection gratings in large TB pulse compression filters
IEEE transactions, MTT-21, no 4, pp 195-205, Apr 1973

47. Williamson R C
Reflection grating filters
SAW filters, Matthews (Ed), Wiley-interscience

48. Woods R C
Dispersive delay lines using 180 deg reflecting metal dot arrays
1982 IEEE ultrasonics symposium proceedings

49. Wright P V, Haus H A
A closed-form analysis of reflective array gratings
1980 IEEE ultrasonics symposium proceedings, pg 282ff