The Bell System Technical Journal

Vol. XXIII *January, 1944* *No. 1*

The Discernibility of Changes in Program Band Width*

By D. K. GANNETT and IDEN KERNEY

One of the factors that should be considered in determining how wide a transmission band is required for high fidelity broadcasting is the ability of people to perceive the effects of restricting the band to various limits, when listening to typical radio programs. Tests are described in which this was directly measured. The tests were concerned only with the physical ability to hear the differences in band width and disregarded the question of the enjoyment or aesthetic appreciation of wider bands. It is concluded that changes in band width are detectable about twice as readily with music as with speech; that one must go from 8 to 15 kc. to obtain a change as readily detected as a change from 5 to 8 kc.; and that both these changes, for speech, are just sufficient to have an even chance of being detected by listeners having experience in such tests.

THE question of how wide a frequency band it is necessary to transmit to provide high fidelity broadcasting involves consideration of a number of factors. Among these are the limits of hearing of the human ear, the spectra of program material, the aesthetic sensibilities of listeners, the effect of room noise in studios and homes, and the acoustic properties of rooms. A true engineering solution of the problem would attempt to assign numerical values to each of these factors, and then to combine them in some way to obtain a figure of merit versus band width. Sufficient information to do this in a complete and satisfactory manner is not available, however, and in practice the final answer is usually obtained by the exercise of judgment, bolstered by such technical data as can be found on the component factors.

The first two of the above factors, the limits of hearing and the spectra of program material, have been separately investigated and the results published in the technical literature by a number of experimenters. Because of the intangibles involved, however, even these two sets of data cannot readily be combined, forgetting the other factors, with complete assurance that their contribution to the answer is established. The authors, therefore, undertook a series of tests to measure directly their combined effect.

* This paper is a publication, substantially without change, of a report prepared some time ago before work non-productive to the war effort was suspended.

These experiments tested the ability of critical listeners to hear changes in band width on direct comparison when listening to representative program material. The purpose of this paper is to present the data from these tests. Similar experiments have of course been done before. The excuse for this paper is that the experiments represent a complete set of data and the analysis of the data is believed to be in such form as to be useful in further consideration of the requirements of program fidelity.

The circuit arrangements used for the tests are shown schematically in Fig. 1. The essential features are a source of program, a switch for connecting into the circuit either of two low-pass filters, and a high-quality loudspeaker. Controls for adjusting levels, volume indicators, etc., are omitted from the diagram. The arrangements included a signal visible to the

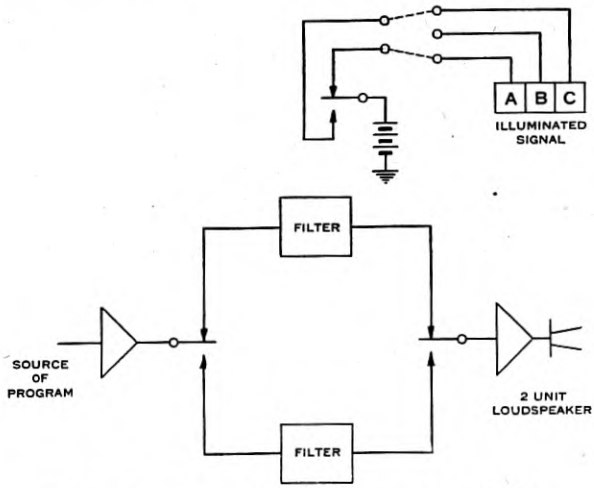

Fig. 1—Arrangement for testing program band widths.

listeners in which one of the letters, A, B, or C, could be illuminated. On a given test two of these letters were associated with the switch so that one letter was illuminated for one position and the other letter for the other position. The choice of letters among the three was varied more or less at random for different tests. Low pass filters were available to provide cut-offs of 3, 5, 8, 11 and 13 kc. When no filter was inserted the band was considered to extend to 15 kc. as this was about the upper limit of transmission of the testing circuits and loud speaker. The lower limit of the transmitted band for all conditions was approximately 40 cycles.

In conducting a test, a group of observers listened to comparisons between two of the available band widths, the conditions being switched every few seconds until a sufficient number of comparisons had been made. The

conditions were unknown to the observers, being designated to them only by the letters in the signal. At the conclusion of the test the observers were asked to mark on a ballot which letter appeared to coincide with the wider band (not which they preferred). A series of tests consisted of comparisons between substantially all of the possible band widths among those available. There were also included in some of the series as a check, one or two tests in which the band width was the same for both positions of the switch. Ten complete series of tests were carried out, two on each of five different programs.

The programs consisted of a dance orchestra, two large symphony orchestras, speech from a male speaker repeating a test sentence, and a radio dramatic sketch. The programs, except for the spoken test sentences, were obtained by special arrangement over direct wire lines from the studio or theater in which the performance took place. The entire system from microphones to and including the loud-speaker had a substantially flat transmission characteristic from 40 to 15,000 cycles, with no filters in the circuit. The loud-speaker was of the two-unit type and was one of a number built for the demonstration of auditory perspective in 1933. The tests were conducted in the program laboratory of the Bell Telephone Laboratories where the acoustic noise level was about +30 decibels. The noise contributed by the electrical parts of the system was considerably below the acoustic noise. The loudness of the programs was adjusted to about unity reproduction, that is, to the volume that would be heard by listeners in a favorable position at the original performance.

The observers were engineers having a considerable experience in tests of program quality. They were doubtless therefore considerably more critical than the average radio listener. The number of observers varied somewhat during the tests but averaged about sixteen. The ages of the observers were in the 30's and 40's so that neither very young nor very old ears were represented.

The immediate outcome of the tests was some 2,000 ballots which were meaningless until analyzed. Before the analysis could be made, however, it was necessary to decide how to express the results.

There are no familiar units to express fidelity or program quality. It was decided therefore to employ the very useful concept of the limen and the liminal unit. These terms have occasionally been applied to other subjective data and may be roughly defined as the least change in a quantity which is detectable. In the present case, if the band widths being compared differ greatly, there will be a nearly unanimous agreement among the observers as to which is the wider. If they differ only slightly, however, many of the observers will vote wrongly for the narrower band and on successive repetitions of the test many will reverse themselves. An average of

a large number of votes will show a plurality for the wider band, the margin of choice increasing as the difference in band width is made greater. A significant measure of the detectable difference in band width will be taken to be that difference such that 75% of the observers correctly select the wider band and 25% wrongly select the narrower band. This difference in band widths will be designated one "difference limen." The sensory effect of a change of one difference limen will be called one "liminal unit".

The significance of the vote of 75 to 25% is assumed to be as follows: On a particular test some of the observers can detect the difference between the conditions while the remainder will guess. Of the latter, half are likely to guess right and half wrong. When 25% vote wrongly they are assumed to be guessing and must be paired with another 25% who also guessed but happened to guess right. Therefore a vote of 75 to 25% is taken to indicate that 50% of the observers were guessing and the remainder could actually detect the difference. The difference limen may now be more specifically defined as that difference in band widths which is detectable to half the observers.

It may be commented that this attempt to explain the definition of "liminal unit" is perhaps over-simple. The observers themselves are frequently uncertain whether they are guessing or are influenced in their choice by some minute difference. The test could be done with a single observer, repeated many times to obtain the same number of observations as with a group. When the conditions are nearly equal he will vote about as often one way as the other, but as the difference between the conditions is increased he will vote a larger per cent of the time correctly for the wider band, just as did the group. When the two conditions are separated by one difference limen he will vote correctly 75% of the time and wrongly 25% of the time, which may be said, in line with the argument given earlier, to indicate that he is guessing half the time and can discern the difference half the time. The difference limen could therefore be defined as that threshold difference for which there is an even chance of its discernment by a listener.

Having chosen a method of expressing the results, the analysis can now be attacked. The first step is to group together all tests on similar types of program material, and to determine for each band width comparison the per cent of votes for the wider and narrower band, respectively. The data thus obtained for music and speech are shown by the solid curves of Figs. 2 and 3. A curve labeled 8 kc., for example, shows the per cent of the total votes which selected as the wider each of the other band widths to which 8 kc. was compared. The points, although somewhat irregular, fell systematically enough to permit drawing the smooth curves with the application of some judgment and having due regard to the necessary symmetry between them. (For example, the 8 kc. curve at an abscissa of 5 kc. must

agree with the 5 kc. curve at an abscissa of 8 kc.) A much larger volume of data would be needed to obtain points falling accurately on a smooth curve. To facilitate obtaining the best approximations, the curves were plotted on several kinds of coordinates, including rectangular, semi-logarithmic (shown in the illustrations), probability and logarithmic probability.

The dotted curves were interpolated between the solid curves and progress in steps of 1 kc. The interpolation was readily accomplished with consider-

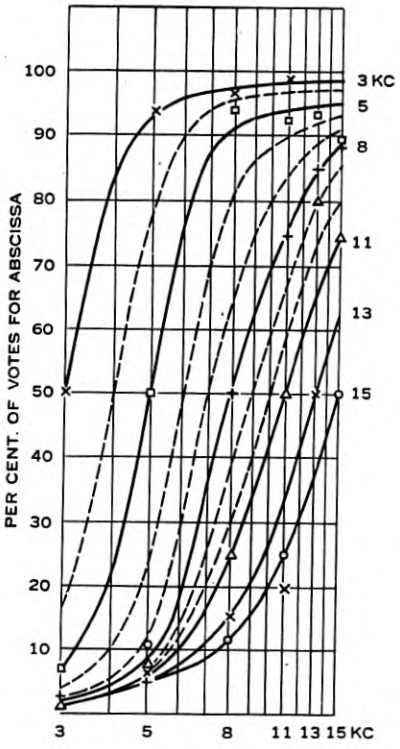

Fig. 2—Music

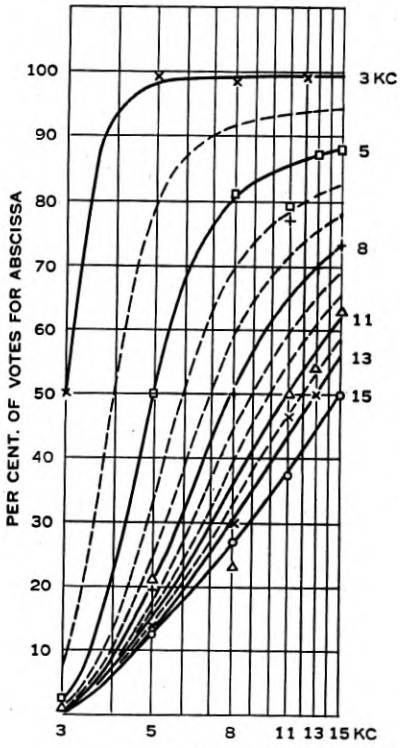

Fig. 3—Speech

Figs. 2 & 3—Detectability of changes in band width.

able accuracy. For example, points for the 10 kc. curve are obtained from the values of each of the solid curves corresponding to an abscissa of 10 kc.

From these curves, the difference limens for each band width were determined by reading directly the bands corresponding to votes of 25% and 75%. The bands at which these votes occur therefore by definition differ from the reference band by one limen. The following table gives the intervals of one limen as thus derived from the curves.

DIFFERENCES IN UPPER LIMIT OF PROGRAM BAND IN KC, CORRESPONDING TO ONE LIMEN

Music	Speech
3— 3.6	3— 3.3
3.3— 4— 4.8	3.4— 4— 4.8
4.1— 5— 6	4.1— 5— 6.9
5 — 6— 7.4	4.6— 6— 9.4
5.8— 7— 9.3	5.1— 7—12.8
6.4— 8—11	5.5— 8
6.9— 9—12.2	5.8— 9
7.4—10—13.4	6.2—10
8 —11—15	6.4—11
9.8—13	7 —13
11 —15	7.6—15

The difference limens are seen to vary with the frequency of cut-off, increasing as the frequency increases. Since each difference limen corresponds to a sensory effect of one liminal unit, it is obvious that the reciprocal

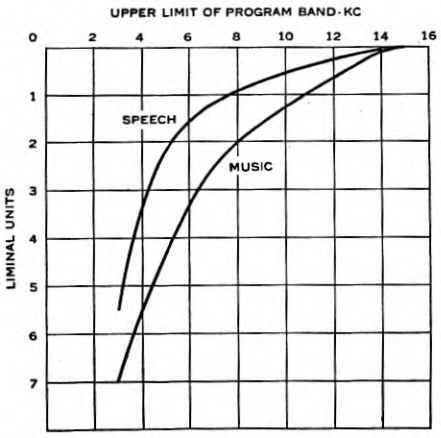

Fig. 4—Ability to detect changes in program band width.

of the difference limen gives the rate of change of liminal units with changes of program band width in terms of liminal units per kilocycle. Therefore, curves of liminal units versus the upper limit of the program band may be constructed from the figures in the table. Such curves are plotted in Fig. 4. The actual mechanics of the process used to plot the curves was as follows, taking the data for "music" for illustration. The lowest frequency occurring in the table is 3 kc., and it is seen that raising the band width to 3.6 kc. will bring about a subjective increase of one liminal unit. Therefore, on an arbitrary scale, 3 kc. was plotted at 0 and 3.6 kc. at one liminal unit. Next a smooth curve was drawn through these points and the location of 3.3 kc. (next line of table) was determined by interpolation. Since 4 kc. is one liminal unit above 3.3 kc., and 4.8 is one liminal un't above 4 kc., these points were plotted and the curve extended through them. By a similar process

the curve was extended step by step up to 15 kc. Finally, the origin was shifted so as to express the liminal curve with respect to 15 kc. instead of 3 kc.

It was mentioned above that a number of tests were introduced without the knowledge of the observers in which the conditions were not changed, the band width remaining constant while the illuminated letters were switched. This produced the most interesting psychological result that observers voted nearly two to one for the letter appearing in the right-hand position in the signal, on each of the six tests of this kind. This raises the question as to whether this effect impaired the results on the other tests.

In the course of the tests, comparisons between each pair of band widths were presented 10 times, 6 times with music and 4 times with speech. The letters corresponding to the two conditions were assigned more or less at random from the three letters A, B, and C. Taking 11 of these groups of tests in which the narrower band was represented about as often by the right hand as by the left hand of the pair of letters chosen, the average vote for the right-hand letter was 51.1% and for the left-hand letter was 48.9%. The difference between these two figures is too small to be significant. It is therefore concluded that when there was a real difference, the observers were not measurably influenced by their slight subconscious predilection for the right-hand letter. It would be interesting to correlate this phenomenon with the right or left-handedness of the observers. This point illustrates the extreme care that must be taken in conducting judgment tests of this sort to insure that no irrelevant factors affect the statistical result.

The curves of Fig. 4 permit drawing the following conclusions:

1. Increases in band width can be detected up to 15 kc. for both music and speech. The fact that this is true for speech is rather surprising. However, above about 5 kc., changes in band width are twice as readily detectable on music as on speech.
2. It requires an increase in band width from 8 to 15 kc. to be as readily detected as an increase from 5 to 8 kc., for both speech and music.
3. The following intervals correspond to one liminal unit and are therefore just discernible half of the time to the observers:
   Speech: 5 to 8 kc.; 8 to 15 kc.
   Music: 5 to $6\frac{1}{2}$ kc.; $6\frac{1}{2}$ to 8 kc.; 8 to 11 kc.; 11 to 15 kc.

In considering these conclusions, the fundamental assumption and limitations of the data should be borne in mind. First, the data were obtained from tests with a certain group of observers and on certain program material. Curves of somewhat different slope would doubtless be obtained with observers of different average age, experience, musical appreciation, etc. It is likely, however, that this would affect the absolute importance of the different intervals in liminal units rather than the relative values. As noted

earlier, the observers in these tests were considerably more experienced and critical than average radio audiences. The program material tested was representative of most of the programs on the air, but different results would be obtained with material markedly different in nature. This would probably be particularly true of selected sound effects. Secondly, it should not be forgotten that the results are based only on the ability of the ear to detect the changes, with no weighting for factors such as aesthetic values or per-

TABLE I

	Upper Frequency Limit Versus Unrestricted Band, Corresponding to One Liminal Unit
Musical Instruments	
1. Flute	13,500 cycles
2. Snare Drum	13,000
3. Violin	13,000
4. Soprano Saxophone	12,700
5. Oboe	12,700
6. 14 in. Cymbals	12,000
7. Bass Clarinet	10,500
9. Piccolo	10,200
9. Bassoon	10,000
10. Cello	9,800
11. Bass Saxophone	8,600
12. Clarinet	8,500
13. Trumpet	8,300
14. Bass Viol	7,800
15. Trombone	7,200
16. Bass Tuba	6,300
17. French Horn	6,100
18. Piano	5,600
19. Bass Drum	4,300
20. Timpani	3,500
Speech	
Male	7,300
Female	9,200
Sound Effects	
Footsteps	12,000
Handclapping	15,000
Key Jingling	15,000

sonal preferences, or for the effects of room noise and other factors present in the practical case. Thirdly, it should be appreciated that comparison tests such as these are very sensitive tests, showing up differences that could not be detected under usual home listening conditions.

It is of interest to compare the above results with previously published data. In a paper "Audible Frequency Ranges of Music, Speech and Noise,"[1] W. B. Snow gave data for 20 musical instruments, certain noises, and

[1] Jour. Acous. Soc. Amer., July 1931; *Bell Sys. Tech. Jour.*, Oct. 1931.

speech. The data showed the frequency limitations as compared with unlimited bands (about 15 kc.) which yielded a vote of 60 to 40%, and 80 to 20% among a considerable number of observations. In Table I these data have been interpolated to determine the limits that would correspond to a vote of 75 to 25%, in line with the criterion assumed in this paper. In making the interpolation, it was assumed that the curve of per cent of observers voting correctly for the wider band versus logarithm of the frequency is a straight line in the range of interest.

TABLE II

	Lower Frequency Limit Versus Unrestricted Band, Corresponding to One Liminal Unit
Musical Instruments	
1. Bass Viol	53 cycles
2. Bass Tuba	55
3. Timpani	60
4. Bass Drum	72
5. Bass Saxophone	72
6. Bassoon	74
7. Bass Clarinet	80
8. Cello	83
9. Snare Drum	87
10. Piano	95
11. Trombone	110
12. French Horn	125
13. Clarinet	140
14. Trumpet	160
15. Soprano Saxophone	210
16. Violin	230
17. Oboe	240
18. Flute	250
19. 14 in. Cymbals	370
20. Piccolo	510
Speech	
Male	115
Female	190
Sound Effects	
Footsteps	95
Handclapping	135
Key Jingling	915

It is difficult to interpret these data from individual instruments in terms of results to be expected from whole orchestras and other music as usually heard. However, comparing Table I with Fig. 4, it will be seen that the frequency limit determined from the present tests as corresponding to one liminal unit for music falls about one third the way down the list of instruments in the table, and the limit corresponding to two liminal units falls about two thirds down the table, which seems reasonable. Also the frequency limit found in the present tests to correspond to one liminal unit for

speech lies between the figures given in the table for male and female speech, which is a good check.

The present tests did not include measurements on the lower end of the frequency band. However, some clue to the results that would be expected may be obtained from Mr. Snow's paper. Table II, derived from Mr. Snow's data in a manner similar to that just described, gives the lower limit of the frequency band corresponding to a degradation of one liminal unit compared with transmitting a much lower frequency.

The frequency corresponding to one liminal unit for speech may be taken as the mean of the figures for male and female speech, or about 150 cycles. In the case of music, it may be expected that at the lower as well as the upper end of the frequency range one liminal unit for an orchestra should fall about one third the way down the list of individual instruments, and two liminal units about two thirds the way down the list. This would make one liminal unit for music correspond to about 80 cycles and two liminal units to about 150 cycles. This speculation leads to the interesting hypothesis that the relations are probably the same at the lower as at the upper end of the frequency scale, that is, changes in band widths are twice as readily detected for music as for speech, and that the frequency limit corresponding to one liminal unit for speech corresponds to two liminal units for music.

CHAPTER V

Use of the Etch Technique for Determining Orientation and Twinning in Quartz Crystals

By G. W. WILLARD

This paper is one of a series of papers dealing with piezoelectric circuit elements and their manufacture.[1] Certain parts of the paper are not new or original, but have been added for the sake of completeness and for the convenience of the reader.

5.1 Introduction

THE manufacture of piezoelectric plates from crystalline material involves orientation problems not encountered in the fabrication of objects from non-crystalline materials. The reason for this is that crystalline materials have physical properties which vary with the orientation, or direction, in which they are measured. Since the operating characteristics (activity, frequency, and temperature-coefficient) of the finished piezoelectric plate depend, not only upon the shape and dimensions of the plate, but upon the physical properties (electrical, elastic and thermal) of the crystalline material, the finished piezoelectric plate must have a *specific orientation* with respect to the material as well as a specific shape and dimensions. In the case of quartz piezoelectric plates the orientation problem is complicated by two factors. First, a large portion of the available natural quartz crystals lack such natural faces as are required to determine accurately the structure-orientation from the shape of the original stone. Thus the raw stones must be examined for structure orientation by physical instruments before even the first cuts may be made. Secondly, a large portion of natural quartz crystals are twinned, i.e. not of the same structure orientation throughout the stone. The boundaries of the respective, homogeneous regions are not predictable, and cannot be completely located in the uncut stone. Thus the processing of quartz involves a step by step examination for twinning boundaries and orientation as the raw stone is cut into sections, the sections cut into bars or slabs, and the bars or slabs cut into blanks. Even when using untwinned stones the orientation must be redetermined and corrected at each cutting step when making such plate types as require very exact orientation.

The most widely used methods of determining the structure orientation

[1] See B.S.T.J., Vol. XXII: No. 2, July 1943 for Chaps. I and II; No. 3, Oct. 1943 for Chaps. III and IV.

of quartz are: (1) by optical effects (birefringence and rotatory power), (2) by X-ray reflections from atomic planes, and (3) by the use of etch pits which are developed when the quartz surface is etched in fluorine compounds. Other methods are or may be used in rather special cases. For example, in finished plates of known orientation types, the electrical axis direction is distinguished from other directions by electrical polarity tests (on tension or compression), or a plate known to be one of several types may be tested in an electric circuit for activity, frequency and temperature-coefficient, to determine which type it is. The selective fracture characteristics of quartz offer another method of determining orientation. Microscopic fractures resulting from grinding a quartz surface may be used for determining orientation. Thus unetched, ground, Z-cut surfaces of quartz give a hexagonal figure, when examined by pinhole illumination, which may be used to determine the approximate orientation (but not sense) of the electric axes.[2]

By optical methods (see Chapter II) it is possible to determine the orientation of a quartz body relative to only one direction of the structure, the optic or Z axis. Thus optical methods are limited to determining the angle between the optic axis and a line or surface of the body (but not the rotation of that line or surface about the optic axis). Twinning of the "optical" variety may be detected optically, even when located internally, but the determination of its location in depth is approximate.

By X-ray methods (see Chapter III) it is possible to determine the structure orientation of a quartz body exactly and completely. However, this method is limited in application by the complexity of analysis, except when the approximate orientation is already known. Though twinning can be detected on the surface of the body, it is not generally feasible to explore the surface to locate twinning boundaries. Further, though positive or negative sense of angular orientation is obtainable by X-rays, this part of the complete determination is not reliable unless the specimen examined is known to be free of twinning, or unless the twinning boundary locations are known. Thus X-ray determinations of orientation are generally limited to determining exact orientations in quartz bodies of approximately known orientation (which includes the case in which only one axis is approximately known).

The etch method of determining orientation is commonly used in conjunction with the optical and X-ray methods to give the information that those methods do not give. The etch method, as most commonly and practically applied, does not give exact orientation angles, nor is it applied to specimens of entirely unknown orientation. However, when a surface of approximately known orientation is etched, it is possible to determine approximately the complete orientation (including sense) of the specimen, and further to detect at this surface both electrical and optical twinning and to

[2] See Fig. 5.20, and further explanation at the end of Sec. 5.53.

determine exactly the twinning boundary locations. The detection of twinning and twinning boundaries by this method has been practiced for years. The determination or orientation and sense of orientation has been exploited only more recently. At present the etch methods play an important and extensive role in the processing of quartz plates, not only in the routine determination of orientation, but also in the detection of twinning so that the most economical cutting methods may be practiced.[3]

5.2 Twinning (General)

Although the problems related to twinning are largely those of determining orientation of the crystal structure, the nature and prevalence of twinning in crystal quartz presents a special group of problems that would be absent were the twinning absent, and hence are separately grouped as twinning problems. As pointed out in Chapter IV, there are only two common types of twinning in the commercial quartz used for piezoelectric plates, namely, electrical and optical twinning. A simplifying feature of both these types is that the structure axes (optic axis and electric axes) of all portions of a single crystal are parallel each to each. However, they are not of the same sense, or handedness. The difference between the two types is as follows:

In a crystal which is only ELECTRICALLY TWINNED, the crystal is entirely of one handedness (either right or left), but one portion is of OPPOSITE ELECTRICAL SENSE to another portion, i.e., the electric axes are of opposite sense.

In a crystal which is only OPTICALLY TWINNED, one portion of the crystal is of OPPOSITE HANDEDNESS, and electrical sense, to another portion. This twinning (but not electrical) is detectable by optical means (polarized light) and is named optical twinning for this reason.

The extent of twinning that may be present in commercial crystals is seen in Fig. 5.1, which shows both electrical and optical twinning boundaries at the top surface of some Z-cut (basal) sections of quartz (which were cut up for the manufacture of quartz oscillators). Though the crystals are seldom entirely free of twinning, they do not on the average run as badly twinned as here shown. These views, taken by means to be described, correspond to what one sees when examining an etched quartz surface by reflection from a strong light.

Since untwinned finished plates must be cut entirely from one twin or another (not across a boundary), and since the proper sense of angular orientation of the plate is opposite for two adjacent electrical twins, the economic utilization of twinned quartz is a difficult problem.[4] It involves cutting the

[3] Etching is also used on finished plates for removing grinding debris, and for frequency adjustment.

[4] As herein used, a *twin* is one of the homogeneous, untwinned portions of a twinned crystal.

stone into separate parts when the twins are large enough to be utilized separately. Further, at some stage before reaching the finished plate all twin portions but one must be cut away.[5]

In this connection it is important to note a size and form difference between electrical and optical twins. Fig. 5.2 shows the appearance of twinning boundaries when only ELECTRICAL TWINNING is present. Note that electrical twins are commonly large, hence may often be separated ap-

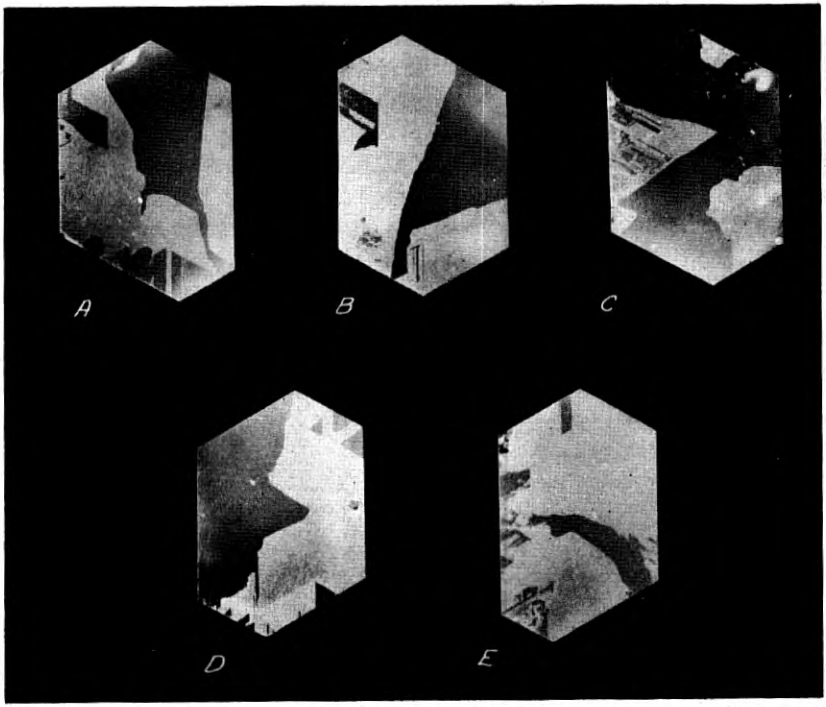

Fig. 5.1—Examples of ELECTRICAL and OPTICAL twinning, as exhibited at the etched surface of Z-cut sections. These examples are typical of an appreciable portion of the quartz that is cut up for quartz plates.

proximately along a boundary and both portions utilized. Fig. 5.3 shows the appearance of twinning boundaries when only OPTICAL TWINNING is present. Since optical twins are commonly small and in the form of thin laminations, it is seldom possible to cut optical twins apart and use both parts separately.

The conventions here used, regarding handedness and axial sense, are

[5] See Section 5.7 for the possibility of utilizing partially twinned finished plates.

according to those of the proposed "I. R. E. Standard."[6] Figure 5.4 shows the relation of these conventions to the natural faces of right and left quartz, to the electric charges developed on compression and tension, and to the more common cuts of oscillator plates. Also given are the relations of handedness to the conoscope and the polariscope means of detecting handedness (Section 2.7, Chap. II describes these instruments). It is important to

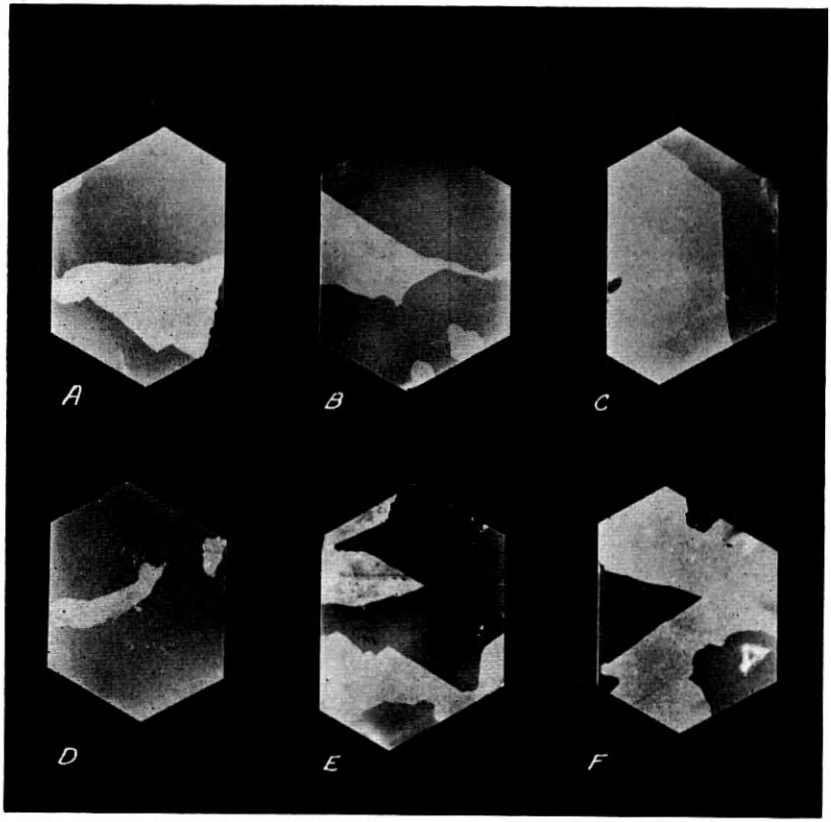

Fig. 5.2—Examples of ELECTRICAL twinning alone. Electrical twins are commonly large, and hence may be cut apart and used individually.

note that AT and CT plates are always cut at such an angular sense, relative to the Z and X axes, as to be roughly parallel to a *minor pyramidal face*, whereas the BT and DT plates are roughly parallel to a *major pyramidal face*. Thus a stone exhibiting these faces may be cut into any of these plates

[6] "Proposed Standard Conventions for Expressing the Elastic and Piezoelectric Properties of Right and Left Quartz", *Proc. I. R. E.*, Nov. 1942, p. 495.

without determining the handedness and electrical sense of the stone (if twinning is negligible). As will be seen later, a similar situation prevails when analyzing etched X-cut sections for cutting into plates.

5.3 Nature of Etch-Pits

When crystal quartz is etched by contact with hydrofluoric acid (or other etching agents) the surface of the quartz is eaten away in such a manner as

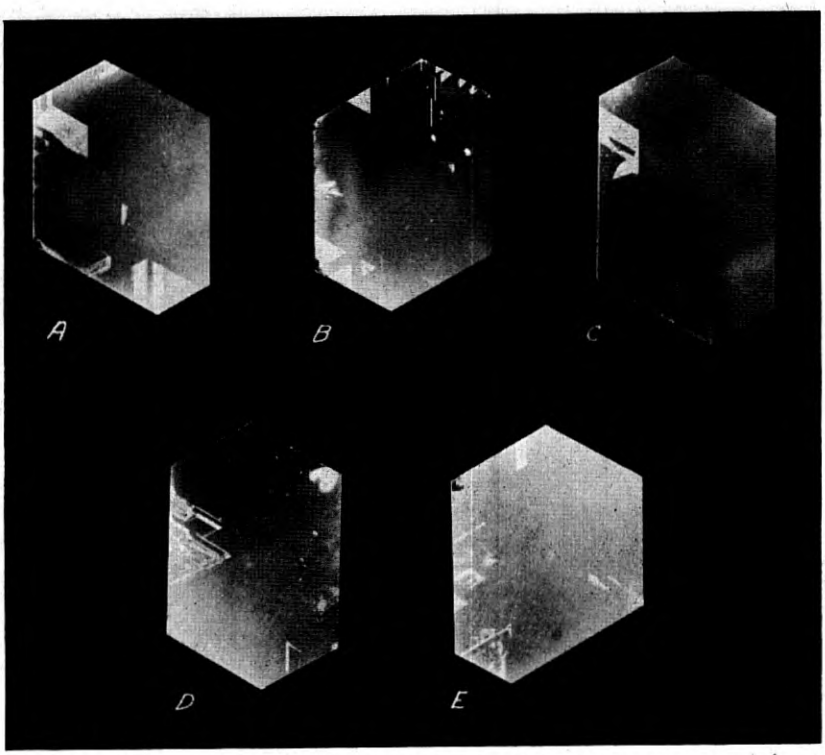

Fig. 5.3—Examples of OPTICAL twinning alone. Optical twins are commonly small and interlayered, and hence may not be separated and used individually.

to leave microscopic *etch-pits* (or hills). These etch-pits are formed of minute facets which are definitely related to the crystal structure. The form of these pits and the orientation of the facets may be used to determine the orientation of the crystal structure at the etched surface being examined.

The general appearance of four types of etch-pits is shown in the photomicrographs of Fig. 5.5. These are the pits that are developed on ground surfaces which are approximately parallel to the well known X-, Y-, and Z-cut surfaces of right hand quartz, by the action of hydrofluoric acid. It is

seen that the positive and negative X-surfaces produce different etch-pits, and are thus usable in determining electrical sense. Further, the pits on all surfaces have directional properties which allow them to be used for determining the approximate directions of the axis which lie in the etched surface. However, to be able to determine orientations from etched surfaces of other

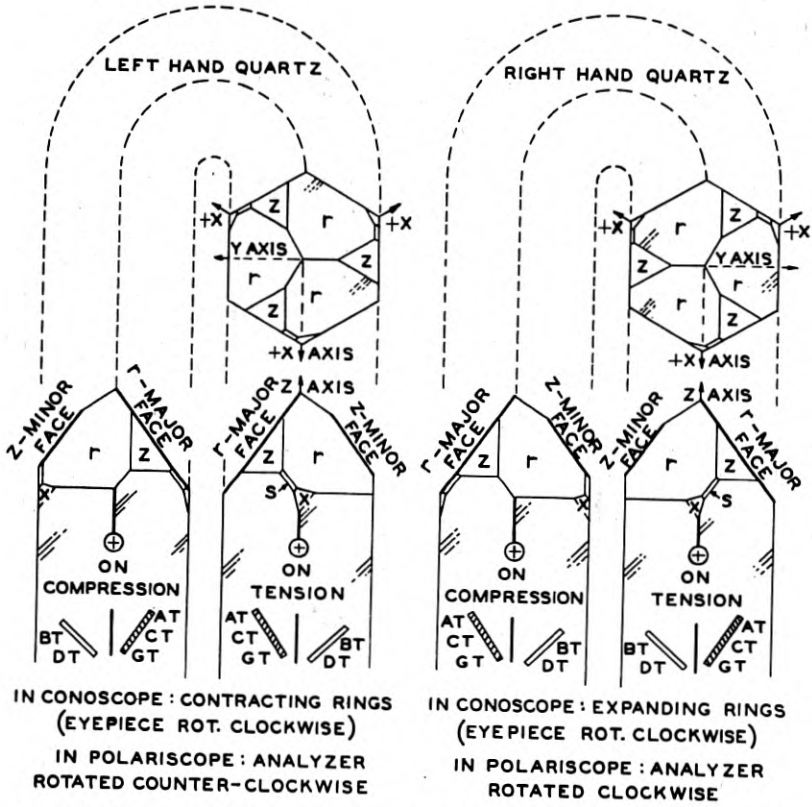

Fig. 5.4—The conventions of handedness, axes, natural faces, and angular sense-of-cut of common oscillator plates, together with the electrical and optical rules for determining these characteristics in unfaced stones.

orientations than those shown above, requires a knowledge of the appearance of the etch-pits developed on such surfaces.

A rather complete catalog of etch-pits on all possible surfaces of quartz was prepared by W. L. Bond,[7] using an etched sphere of quartz (Figs. 5.5, 5.6 are from Bond). Thirty-six different types of etch-pits were obtained and their angular range of coverage was found (the X-, Y-, and Z- surface

[7] "Etch Figures of Quartz," *Z. Kristallogr.* (a) 99, 1938, pp. 488–498.

pits are obtained only on surfaces within 6° to 8°, from the X-, Y-, and Z-surfaces, respectively). Since the development of good etch pits and their exact appearance is considerably affected by the preparation of the surface for etching (fineness of grind), and by the strength of the acid and the length of etching time, and by the manner of illumination when viewing, the

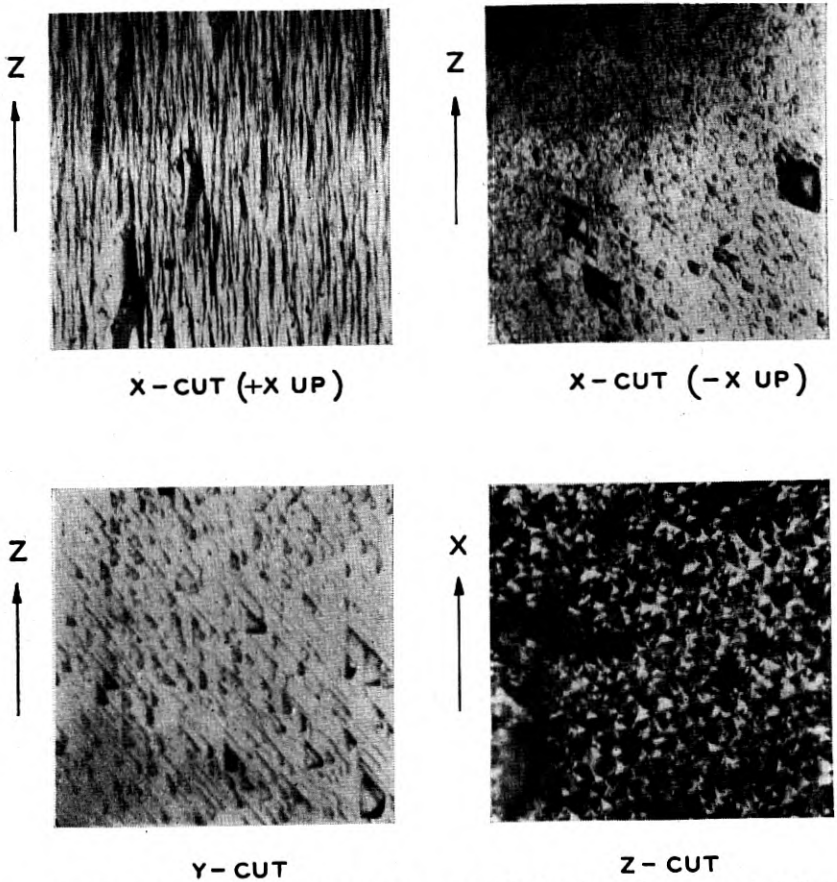

Fig. 5.5—Photomicrographs of etch-pits on the etched surfaces of common orientations. As seen the etch-pits are definitely related to the structure axes of the quartz.

figures shown here do not represent the exact appearance of pits obtained by other manners of development. However, such figures are reproducible.

The use of etch-pits to determine the orientation of a perfectly general surface is complicated by the fact that some different surface orientations give pits not readily distinguished from each other. However, for the surfaces most commonly encountered in quartz plate manufacture the etch-

pits are quite distinctive, when well developed. Use may be made of a microscope or a high powered projector to view the figures. The pit outlines may be aligned with lines ruled on the eye-piece or on the screen, and a fixed marking device may be used to mark the quartz surface with orientation lines. Twinning may be detected by the appearance of different etch-pits as the specimen is moved about. For example, on an electrically twinned X-cut surface both X-cut views of Fig. 5.5 could be found. However, the location and marking of twinning boundaries involves a tedious exploration of the surface, since only a minute portion is viewed at any one time. This exploration may be eliminated if the surface is first viewed by reflection methods where the whole surface and extent of twinning is at once seen (as in Fig. 5.1) and marked.

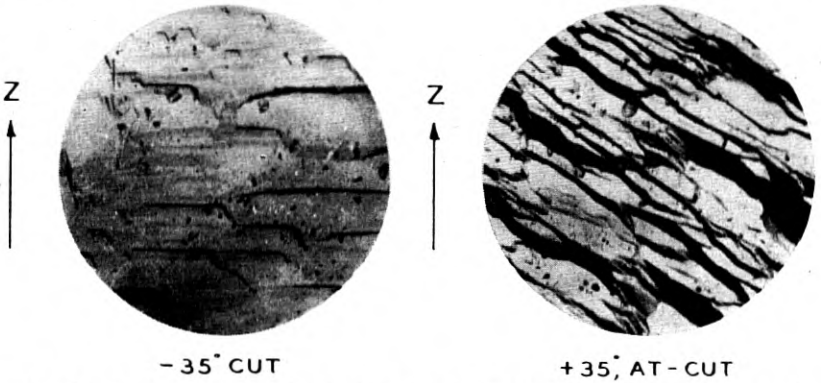

Fig. 5.6—Etch-pits on the etched surface of a +35° AT plate, and on an analogous but wrong sensed −35° plate. This difference in etch-pits may be used in the manufacturing process to determine the right and wrong sensed regions of twinned AT slabs.

A special case where the microscope or projector method might be employed is in the examination of thin AT, BT, CT or DT slabs for twinning and sense of cut. Here the slabs are known to be cut with a reference edge parallel to an electric axis, and with the major faces inclined at 35° to 55° (depending upon the variety of slab) from the optic axis, the sense of the inclination being positive for the AT and CT slabs, and negative for the BT and DT. The effect of electrical twinning on such etched surfaces is shown in Fig. 5.6. The etch-pits of the good +35° AT-portion of the slab are easily distinguished from the analogous −35° (bad) portions. This difference is similarly distinguishable in the other cuts.

Actually, orientation and twinning are seldom analyzed by the method described above, i.e. by examining their appearance in the microscope, or by projection on the screen. The method appears to be far less practical than other methods which depend upon the gross effect, of hundreds of simi-

lar etch-pits, in bending a light beam. By the latter methods the individual etch-pits are never seen, nor does their nature need to be known. Nevertheless, the resultant optical effect of hundreds of similar etch pits is as characteristic of structure orientation as the individual pits themselves.

5.4 Optical Effect of Etch-Pits

The gross optical effect of hundreds of similar etch-pits results from the fact that each of the pits has minute facets which are similarly inclined to those of all the other pits. Though the pits of Figs. 5.5 and 5.6 may not appear to be formed from groups of flat facets they are generally so regarded. "Curved-facets" are theoretically considered to be made up of individual flat-facets which are parallel to possible atomic planes (and hence may be given index numbers as in Chap. III). This view is the same as that taken

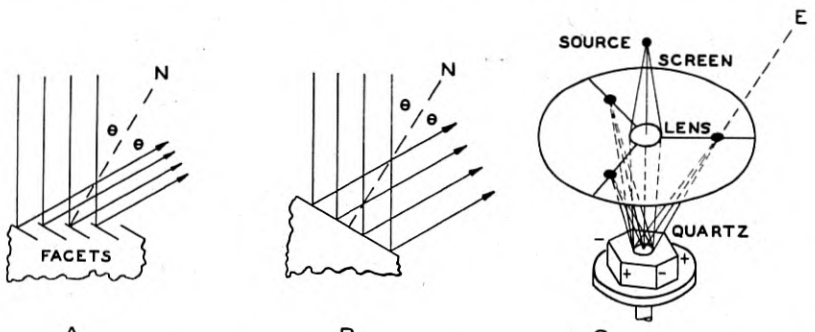

Fig. 5.7—Reflection of light from a single set of similarly oriented etch-pit facets, A, is like that from a single mirror, B. Reflection from all three sets of facets of a Z-cut section will give a three-fold *etch-figure* on a screen, as in C.

with regard to natural faces, which are of course produced by essentially opposite effects, i.e., acid corrosion in the case of etch-pits, and growth from solution in the case of natural faces. Actually, many "curved-facets" give optical effects showing no discernible evidence of individual flat facets. However, the question is academic, so far as use of the pits for orientation purposes is concerned, for such facets are still definitely related to the crystal structure.

Etch-pit facets may be used to *reflect* a light beam into specific patterns or to *refract* the beam on transmission through the material into similar (but not identical) patterns. The different basic optical means of using etch-pit facets are shown in Figs. 5.7, 5.8, 5.9. Included in each figure is a diagram of the effects obtained by illuminating an idealized Z-cut section. This idealized section is assumed to have only simple, equilateral, three-sided

pyramidal etch-pits, oriented relative to the X axes as shown in Fig. 5.5. The actual results obtained with Z sections are more complicated than this and thus indicate that the etch-pits are not exactly as idealized here.

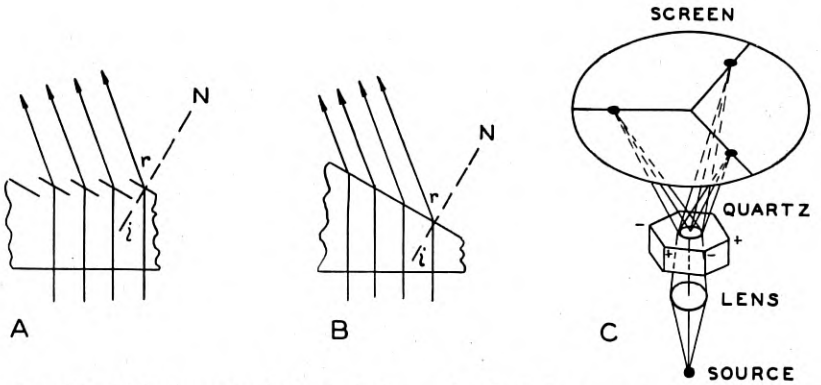

Fig. 5.8—Light transmitted thru a single set of etch-pit facets, A, is refracted as by a prism, B. The three sets of facets of a Z-cut section give a three-fold etch-figure, as in C.

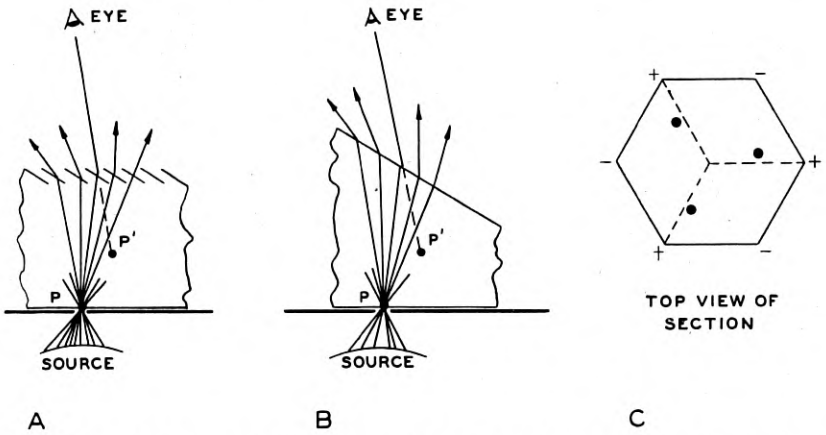

Fig. 5.9—Light transmitted thru a pin-hole is refracted by a single set of facets, A, as it would be by a prism, B. A virtual image of the pin-hole P will be observed at P'. The etch-figure seen down in a Z-cut section is three-fold, as in C.

5.41 The Reflection Method

Figure 5.7 shows the reflection method, where a parallel beam of light striking the etched surface of a Z-section is reflected from one of the three sets of facets as shown in A. Each single facet reflects part of the beam by

ordinary reflection laws, and the whole groups of facets act similarly to a single mirror surface at the same angle, as in B.[8] The individual facets being very minute and of irregular size and spacing, however, cause appreciable diffusion of the beam. The resultant effect of all *three* sets of facets is shown in C, where light passing down through a lens and a hole in the screen is reflected back to three spots on the screen. These three spots are located at equal distances from the incident beam and at 120° intervals around the incident beam. If the quartz section be rotated on its table the spots rotatc around the screen correspondingly. However, lateral motion of the section across the table (without rotation) does not change the position of the spots, if the section be untwinned. If the section is twinned (or more exactly, if the etched surface is twinned) the three-fold figure will shift to a different position (angularly) on crossing a twinning boundary, for the etch pits are oriented differently in the two twins. If the twinning boundary divides the illuminating beam, then both figures appear at once, giving six spots instead of three. It is clear then that twinning, as well as orientation of the section, may be determined from the figure on the screen. The angular relation between the spots and the X-axes of the section will be considered later, where figures of actual sections are shown.

The long used method of examining etched quartz surfaces by simple reflection from a bright light, may also be explained from Fig. 5.7C. If a spot of light on the screen is viewed along the line E, and the screen then removed, the light from the associted etch-pits will fall on to the eye. The illuminated portion of the section will appear bright. If a twinning boundary crosses the illuminating beam and one of the six reflected beams falls on the eye, one of the two illuminated twins will appear bright and the other dark. As the section is rotated, first one twin and then the other will appear bright, and in each case the twinning boundary is sharply defined over the whole region covered by the illuminating beam (the appearance of twinned Z-cut surfaces examined by this means is shown in Figs. 5.1, 5.2, 5.3). Due to the greater complexity of etch-pits than here idealized, the reflected beams are not so sharply defined as to require exact location of the eye relative to the incident beam and the section. Further, when a broad unfocused light source is used, it is possible and convenient to detect twinning boundaries merely by holding the section in the hand and rocking it about in various directions until a brightness contrast is observed. Though the brightness contrast is usually not marked by this simple examination it suffices for many purposes.

[8] That the effect of a group of facets is not identically the same as that of a single mirror, is of more concern where lenses are used for focusing. In this case the displacement of the mirror facets causes a displacement of the focus of the beam from each facet. For beams of small angular range this is of little importance.

5.42 The Transmission Method

Figure 5.8 shows one form of the transmission method of examining Z-cut etched surfaces. A parallel beam of light passing normally up through the bottom polished surface and the top etched surface of a section will be bent by refraction only at the etched surface, as in A. Each facet refracts the light by ordinary laws of refraction, and the whole group acts similarly to a single refracting surface at this angle, as in B.[9] The resultant effect of all three sets of facets is shown in C (where a lens is added for focusing the light beam). If the incident beam is not normal to the bottom surface there is an additional bending of the beam at this surface. If the incident surface is not polished (or rendered optically flat, with a cover glass and immersion fluid, for example) the diffusion at this surface will mask or completely destroy the desired effect.[10]

5.43 The Pinhole Transmission Method

Figure 5.9 shows the pinhole form of the transmission method, as applied to the examination of Z-cut etched surfaces. Here a section with a top, etched surface is illuminated from below through a small hole with a wide angle of illumination. The light radiates upward in all directions from the pinhole, and in passing through the upper etched surface is refracted by a single set of etch facets as in A. With the eye placed above the pinhole (and section), certain of these rays will fall on the eye. The eye then sees a virtual image of the pinhole P displaced to P′, elevated from the level of P, and along the line of the ray which enters the eye. The effect of a group of facets is similar to that of a single prism, as in B.[11] The resultant effect of all three sets of facets of a Z-cut section is shown in C, where the section is viewed from directly above and no optical system is shown. Only the three virtual images of the pinhole are seen and they are located down in the quartz (roughly two-thirds of the way down).

Though the desired effect is due entirely to the top, etched surface, the nature of the bottom surface may cause a deleterious masking effect, which must be considered in the design of an instrument. Due to the diffusing effect of irregularities in the top surface it may act somewhat as a screen upon which the extended light source shown in Fig. 5.9A, B may be imaged by the pinhole. This extraneous image occurs if the bottom surface is polished, and to some extent if the surface is semi-polished, strongly etched, or oily.

[9] See footnote 8.

[10] Similar optics hold if the section is illuminated from the etched side instead of the polished side.

[11] See footnote 8.

This difficulty may be entirely obviated by the introduction of a diffusion screen directly adjacent to the pinhole.[12]

It might be noted that if it be desired to project or photograph the pinhole figure, one must focus on the virtual image which lies between the top and bottom surfaces of the etched specimen. In the simple case diagrammed in Fig. 5.10, it is assumed that the camera lens is at a distance from the section and directly over the section, so that the rays to the lens are essentially normal to the section. For a section of thickness T, and index of refraction n, the elevation E of the virtual image from the bottom surface of the section is given by: $E/T = 1 - \sqrt{1 + R^2/T^2}/n$. Here R is the radial displacement of the virtual image from the axis of the pinhole and is readily observed and measured. Also, R may be calculated from the thickness of the quartz T, the angle θ between the facets and the gross surface, and the in-

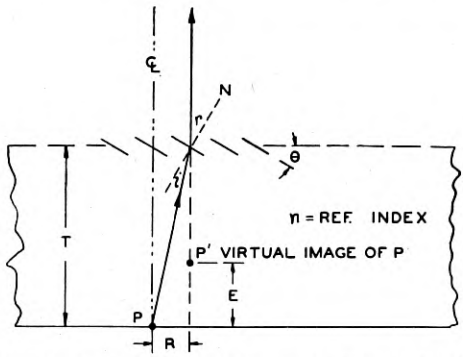

Fig. 5.10—The elevation E of the virtual image may be calculated from the thickness of the etched section T, the radial displacement of the image R, and the index of refraction n; or from T, n, and θ, the angle between the facets and the gross surface.

dex n, (or θ may be calculated from T, R, n) by: $R/T = \tan(\theta - \sin^{-1}[(\sin^{-1}\theta)/n])$. Commonly, pinhole figures from quartz which is weakly to moderately etched (up to one hour in concentrated HF) have a maximum diameter (or double radial displacement) $2R$, nearly equal to the thickness of the section. Since the elevation of the image, E, depends upon its displacement R, an extended virtual image is not in a single plane and cannot be exactly focused (the elevation is commonly about one-fourth to one-third of the thickness of the section). The diameter of the pin-hole must always be kept small compared to the thickness of the section to give sharp figures (and the length of the pinhole must be small compared to its diameter).

[12] The diffusion screen may be a sheet of white paper placed over the pinhole, or a piece of flashed glass placed under the pinhole, with the flashed side against the pinhole. In either case it is usually necessary to increase the light intensity by focusing a concentrated light source onto the pinhole with a lens.

Choice of one of the four above methods of examining etched surfaces for twinning and orientation, depends upon many factors, as will be noted in the following section. The pinhole method is used wherever possible because of the simplicity of the optical system and the brilliance of the figures obtained.

5.5 Etch-Figure Instruments

Herein are described several instruments which have been designed for shop use in determining orientation and twinning of etched quartz sections and slabs. Their basic principles of operation are as described above. The nomenclature of handedness, sense of axes, sense of cuts, natural faces, etc. is according to Fig. 5.4, as explained at the end of Section 5.2.

The etch-figures and reflection patterns obtained on these instruments vary with the preparation of the specimen (i.e. the type of grind and the type of etch). A complete study of these factors would include a variation of the grind from a very coarse grind to polishing (and include saw-cut surface), and a variation of the etching time from short to very long, and the strength and kind of etching agent. Here chosen for illustration are the simplest practical preparations, namely, the coarsest grind usable, and the shortest etching time (in hydrofluoric acid). The etch-figures are thus markedly different than some which have appeared in the literature. Further, the photographic reproduction of etch-figures on paper, is not exact due to the limited contrast range of the paper. Thus in the accompanying illustrations detail is lost in the brilliant portions of the etch-figures in order to show details in the weaker portions, and vice-versa.[13]

5.51 The Reflection Oriascope

Fig. 5.11 shows diagrammatically a reflection "Oriascope", which may be used on specimens with a single flat etched surface. By the reflection principle of Section 5.41 figures are obtained on a viewing screen. Due to the relatively weak figures obtained by reflection from weakly etched surfaces, the viewing screen must be enclosed in a well blackened enclosure, and viewed through an eye chute. The screen is ruled with appropriate lines, relative to which the figure is aligned by turning the specimen on the table. The table is mounted so that when the specimen is properly oriented, the table may be slid to the right or left over a marking template, and marked through the template with appropriate lines to indicate the desired axial orientations of the specimen.

When used with Z-cut sections it is necessary to have two marking templates, one for each handedness of the quartz, since the three-fold figures

[13] Apparent shifts in etch-figure orientation, with etching time for example, are not to be considered as resulting from an orientation shift of the individual etch-pit-*facets*, but as a shift in the relative areas of differently oriented facets. See Figs. 5.12 and 5.17.

obtained are not aligned with the electric axes of the specimen. They are shifted approximately 12° therefrom, and in opposite directions for the right and left varieties. Figure 5.11 shows a section of right quartz so positioned on the sliding table that the etch-figure therefrom will be properly aligned with three radial lines of the viewing screen. The section need not have natural faces as here shown. With the section so positioned the sliding table is moved over the right-hand marking template, and the section is marked with three radial lines. These lines on the section then give the approximate direction (within 5°) and the sense of the three electric axes of the quartz, positive X-outward. With left quartz the etch-figure is still aligned with the same lines on the viewing screen, but the section is marked through the left-hand marking template (the marking having the same meaning as be-

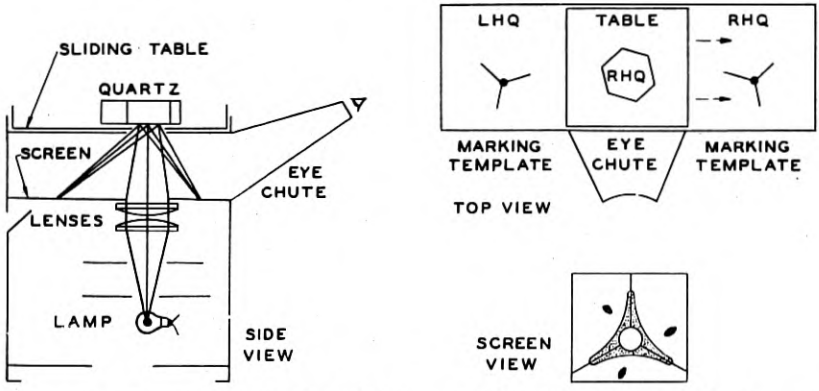

Fig. 5.11—The reflection ORIASCOPE as applied to determining the direction and sense of the X (electric) axes in Z-cut sections. After the etch-figure is aligned on the screen the table and sections are moved over a marking template and the section marked from below with axes.

fore). The section so marked is ready for laying out the approximate cutting directions, the sense of which may be found from Fig. 5.4. The exact cutting directions are obtained by X-rays. It might be noted that ordinarily the handedness of the section is determined in the conoscope (see Section 2.7, Chap. II) before examination on the oriascope. Also the twinning boundaries are previously determined by examination of the etched surface in a spot-light beam.

Figure 5.12A, B show the type of etch-figures obtained on Z-cut sections (in each case the figure is properly aligned with the rulings on the viewing screen). The simpler etch-figure A is obtained on a fine ground (400 carborundum) surface by a weak etch (about 10 minutes in 50% HF). Though the three faint spots, about 40° clockwise from the rulings (for the left-hand

quartz of A) may be used for determining the handedness of the section, it is usually considered more reliable to use the conoscope for handedness determination. The counter-clockwise rotation of these spots in B indicates right-hand quartz. The more complicated etch-figure B, results from etching a fine ground surface too long,[14] or from using a coarse instead of a fine grind. With such figures it is difficult to know which portion of the figure is to be aligned with the screen rulings. Hence the sections must be fine ground and the etching time closely controlled.

The obvious disadvantages of the reflection oriascope (the necessity of predetermining handedness and twinning, and the requirements of fine ground surfaces and closely controlled etching time) are largely overcome by the pin-hole oriascope, later described. However, the reflection oriascope is an

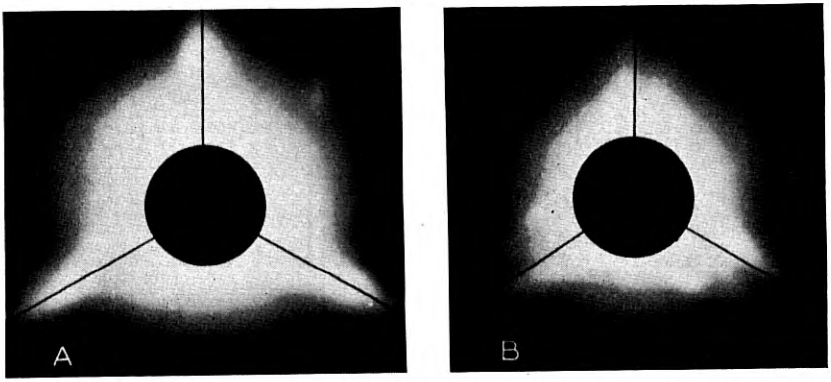

Fig. 5.12—Etch-figures obtained on the reflection oriascope with Z-cut sections (reduced from 11 inches square). A is a good usable figure while B is difficult to use due to its complexity.

excellent explanatory instrument for obtaining experimental etch-figures from surfaces of any orientation, preliminary to devising a special instrument to most advantageously utilize the reflection characteristics found. This fact results from the large and symmetrical screen coverage, and from the fact that only one etch surface is encountered by the light beam (thickness and back surface shape is of no concern).

5.52 The Reflection Twinoriascope

Figure 5.13 shows diagrammatically a reflection "Twinoriascope" designed especially for shop use in detecting and marking twinning boundaries and the sense of orientation in etched AT, BT, CT and DT slabs. When, for ex-

[14] It appears that excessively strong etches (hours long) again give a simple, strong, and reliable figure.

ample, CT slabs are to be examined the tiltable mounting-table is clamped in the 38° position, and the slab placed crosswise on the table (X-axis normal to line of sight, and beveled edge as shown). Upon moving the viewing screen to position *1*, only lamp *1* is lighted, and the slab is viewed by reflected light at a preferred angle. If the slab be twinned, one portion of the slab will exhibit a bright sheen while the other portion is dull by contrast, see two examples in Fig. 5.14, Test *1*. The twinning boundary is now penciled in. The viewing screen is then shifted to position *2* which lights only lamp *2*, and the crystal moved to right or left so that only one twin is illuminated. On the screen[15] will be seen an etch-figure similar to one of the four shown in Fig. 5.14, Test *2*. If either of the two positive-cut figures are observed the illuminated portion of the slab is usable, since the CT plate

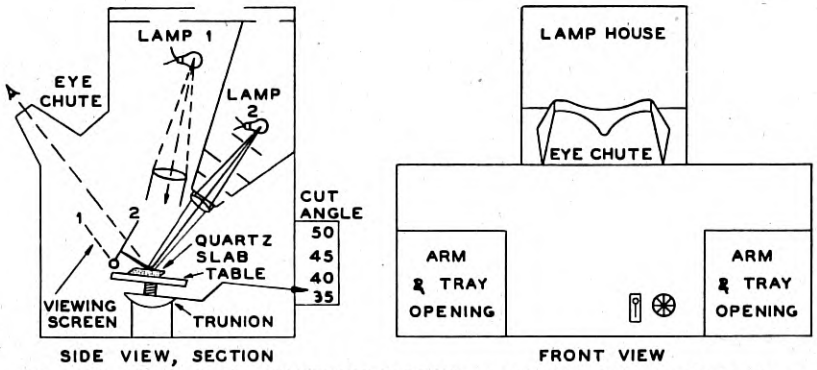

Fig. 5.13—The reflection TWINORIASCOPE for detecting *twinning* (using lamp *1* and no viewing screen, position *1*) and for determining the orientation or *sense-of-cut* (using lamp *2* and the viewing screen in position *2*), of AT- BT-, CT-, or DT-cut slabs. The "cut angle" is set for a CT slab.

must have a positive 38° orientation. The negative-cut, "golf-club", figures are produced by the unusable portion of the plate.

The same procedures are followed with the AT, BT and DT plates, in each case resetting the table to the proper tilt, 35°, 49° and 52°, respectively. The reflection view of Test *1* is the same for all cuts, and the etch-figures of Test *2* are nearly the same (being almost identical for the negative-cut portions of the slabs). However, in the case of AT and CT slabs the *positive-figures* represent *good* portions (since these are positive cuts), and in the case of BT and DT slabs, the *negative-figures* represent *good* portions.

The basic principle of this instrument is as described in section 5.41. As here used, the two optical systems (including the eye and the slab) are so disposed as to obtain the best reflection-contrast in Test *1*, and the most dis-

[15] An excellent screen consists of two sheets of thin sandblasted cellulose acetate.

tinct portion of the etch-figures in Test *2*. That the observations are so similar for this 20° range of cuts indicates that the nature of the etch-pits on these cuts is very similar, (see Fig. 5.6 for the nature of the etch-pits on AT slabs). The angular arrangement of the Test *1* optical system makes use of strongly developed facets which are approximately parallel to the X-axis and inclined at an angle of −57.6° to the Z-axis of the quartz. Within experimental error these facets are parallel to the 01.2 atomic planes and hence are called the 01.2 facets. It is also these facets that give the enlarged

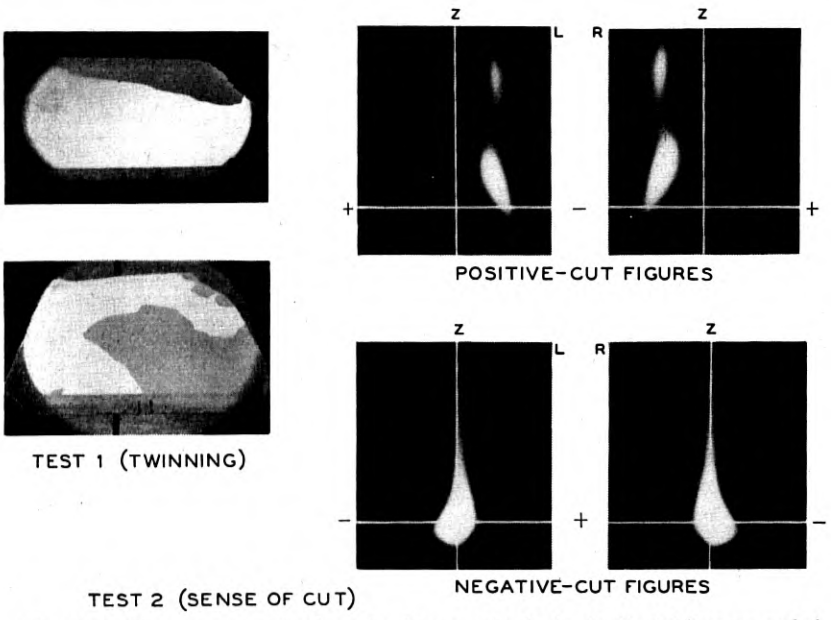

Fig. 5.14—The appearance in the twinoriascope of twinning in Test *1* (two examples) and of the four possible etch-figures in Test *2*. The observance (in Test *2*) of either of the positive-cut figures indicates that the illuminated portion of the slab is a positive cut, while either negative-cut figure indicates a negative cut. These etch-figures for a CT slab, are not markedly different than those for AT, BT, and DT slabs.

head of the golf-club, negative-cut figures. The right and left handedness of quartz results in two figures each for the positive and the negative orientation. Though it is commonly of no interest, it is possible to determine from the etch-figure observed, both the handedness and the electrical sense of the illuminated portion of the slab. The handedness is as indicated by L and R in each etch-figure of Fig. 5.14, and the electric axis is ± to the right or left as indicated by the + and − signs.

Best etch-figures are obtained in the twinoriascope with fine ground (400 carborundun) slabs which have been given a strong etch (40 minutes in 50%

HF). Stronger etching is not deleterious. Very strong etching gives moderately good figures with sawn or coarse ground slabs. For Test *1*, alone, weaker etches would suffice. Under properly controlled conditions of slab preparation and instrument operation Test *2* might be eliminated, for under such conditions the negative-cut portion of the slab is bright, the positive-cut portion is dark. Under shop conditions this means of detecting *sense* of cut appears to be not reliable, especially with *untwinned slabs* (which are either all bright or all dark). The addition of Test *2*, however, gives

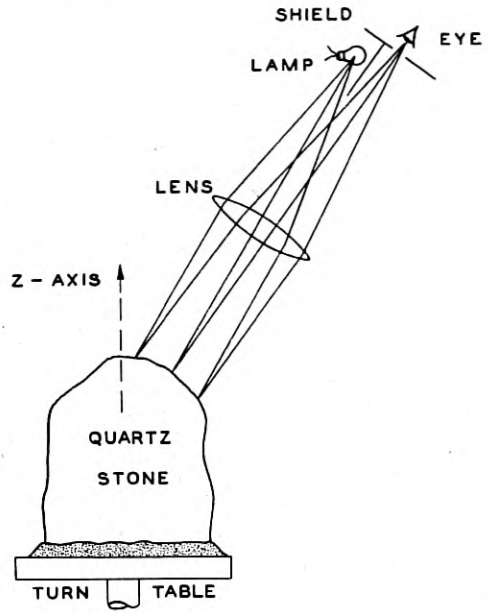

Fig. 5.15—The direction and sense of the electric axes of a sand-blasted and etched raw quartz stone may be determined by reflection of light from the 0.21 facets. These same facets are utilized in Test *1* of the twinoriascope, Figs. 5.13, 5.14.

complete reliability, for if etch-figures are obtained the sense of cut is obvious, if no figures are obtained the slab can be returned for further etching.

The principle of Test *1*, above, has been applied by W. L. Bond to a laboratory instrument for determining the direction and sense of the X-axes in raw quartz stones prepared with a sand-blasted and etched surface. With the stone mounted rotateably about its Z-axis (previously determined by conoscope or inspectoscope), and a light beam properly projected onto the stone, reflection of the light beam to an eye piece or viewing screen will occur whenever the 01.2 facets come into proper angular position, see Fig. 5.15. The approximate direction and sense of the electric axis, or the sense of cuts

to be made from the stone, may be determined from these reflecting positions of the stone, and twinning may be partially explored. Thus if the stone appears to be not badly twinned, it may be cut up at once into slabs of proper sense of cut, without previously sectioning for further examination.

5.53 The Pin-Hole Oriascope

Figure 5.16 shows a "Basic Pin-Hole Oriascope" with auxillary attachments for shop examination of etched Z-cut sections, and Fig. 5.18 the same

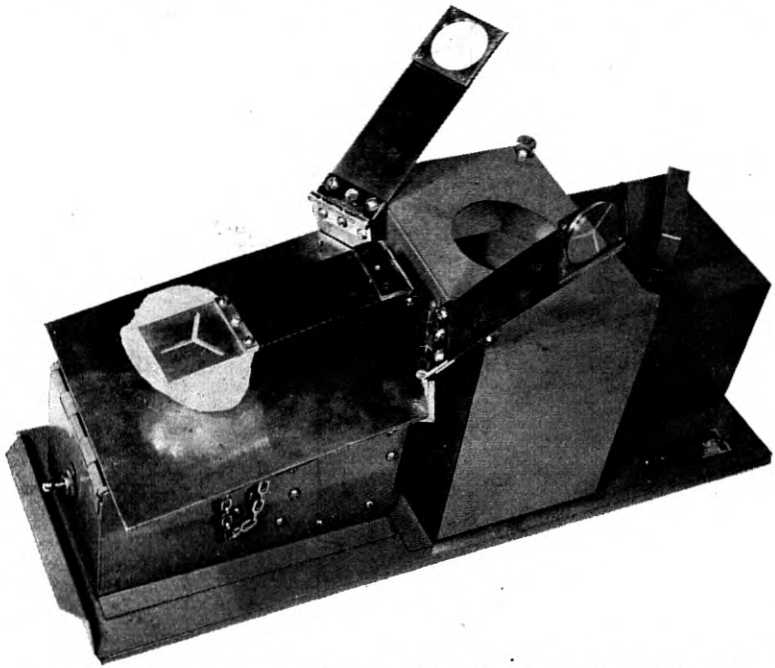

Fig. 5.16—The BASIC PIN-HOLE ORIASCOPE with matching and marking arms for use on Z-cut sections. Twinning, and the direction and sense of the X (electric) axes may be determined and marked on the section.

for X-cut sections. The optical principle of this instrument is according to Section 5.43. Light from a concentrated-filament lamp within the central ventilated housing, is projected horizontally forward by a pair of condenser lenses and reflected upward by a mirror in the forward housing, onto a diffusion-disk placed directly against the pin-hole.[16] The latter is centrally located in the inclined mounting table. Etched quartz sections are placed over this pin-hole and viewed from above. The section may be moved about and examined for twinning boundaries, which are then penciled in.

[16] See footnote 12.

The section is then examined through the ruled window of a matching arm, one of which is shown in use in Fig. 5.18. The section is rotated on the table until the etch-figure seen in the quartz is properly aligned with the lines on the window. Without moving the sections, the viewing arm is replaced with a marking arm, one of which is shown in place in Fig. 5.16. The section is

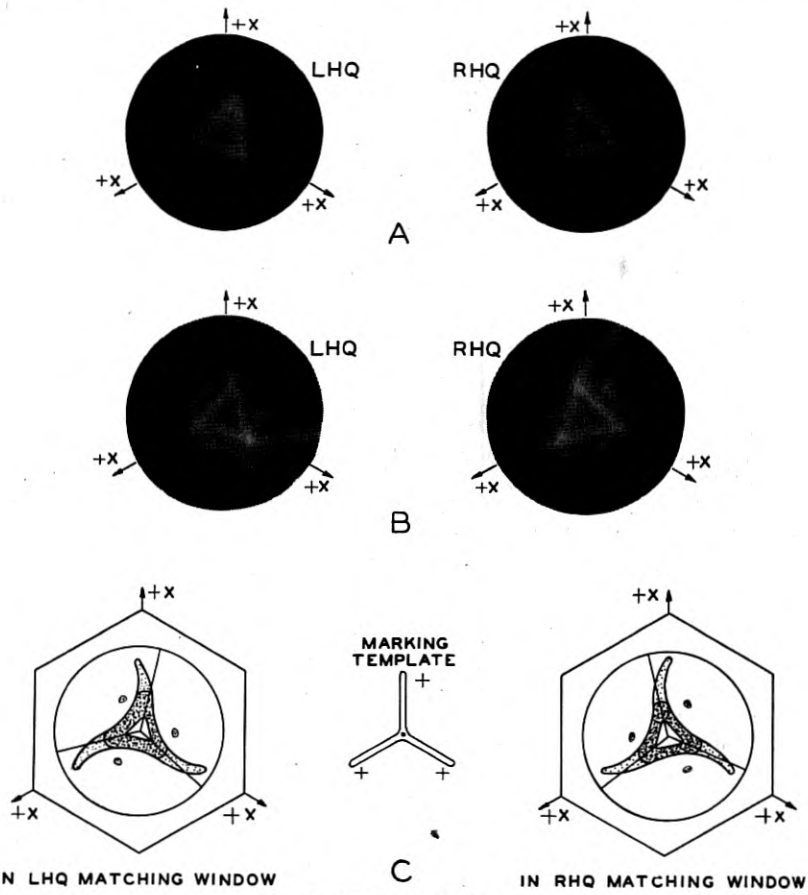

Fig. 5.17—Etch-figures obtained with the pin-hole oriascope in Z-cut sections; A for a fine ground surface and B for a coarse grind. The relation of the etch-figures to the structure orientation of the section is shown in C.

marked through the template of this arm with the desired axes or cutting directions.

Figure 5.17A, B shows the etch-figures obtained with the pin-hole oriascope, on Z-cut sections. Figure 5.17A is for a fine ground surface (600 carborundum) while Fig. 5.17B is for a coarse ground surface (100 carborun-

dum), and in both cases a moderate etch, (20 to 30 minutes in 50% HF). It is noted that the spiralling, outer tails of the etch-figures (as well as other features) denote the handedness of the quartz. Such handedness features are not as marked with fine ground surfaces, nor with weaker etches. The central triangular portion of these figures is used for alignment of the section with the rulings on the marking arm windows. Since this triangular figure is misaligned with the X-axes of the quartz by approximately 12°, and in an opposite sense for the two kinds of handedness, there are provided two matching arms. One is to be used for left quartz and the other for right quartz. The diagram of Fig. 5.17C shows the orientation arrangement of a combination of matching windows and marking template, that results in the section being marked with three radial lines which correspond to the positive X-axes of the quartz. Though this is the most obvious manner of marking Z-cut sections, it is of advantage in practice to obtain a reversed marking on left-hand quartz (by using an oppositely ruled left-hand matching window). By so marking the quartz no further attention need be paid to handedness, see Section 2.4, Chap. II.[17] In either case the relation of the various plate cuts to the axis markings obtained above, may be determined from Fig. 5.4. Since the etch-figures give only approximate orientation X-rays are used for the final determination. That X-rays are not used for the whole determination is as explained in Section 5.1.

With X-cut sections, having a coarse grind (100 carborundum) and a strong etch (30–45 minutes in 50% HF), the etch-figures obtained are like those of Fig. 5.19. Here the positive face of the section gives an entirely different figure than the negative face, as would be expected from the nature of the etch-pits shown in Fig. 5.5. Opposite-handedness gives reversed figures. The four possible figures are oriented with respect to the Z-axis and the major cap face direction of the section "r" as shown in Fig. 5.19A and B. The non-parallelism of the Z-axis and the parallel sides of the etch-figures amounts to three to five degrees. This disposition of figures (relative to quartz axes) is taken into account in the design of the matching and marking arms shown in Fig. 5.18, and diagrammed in Fig. 5.19C. The etched X-cut section is rotated on the mounting table, with the central matching arm in position, until the long straight sides of the "parallelogram" figure, or the long parallel lines of the "H" figure, are parallel to the two parallel-lines ruled on the window of the matching-arm (the parallelogram figure is shown so aligned in C). The figure thus used is compared with the four figures sketched on this matching-arm, to determine which of the two marking arms is to be used for marking (note arrows giving this indication). The proper marking arm is lowered onto the section and used to

[17] The instrument of Fig. 5.16 has a still different arrangement of matching and marking arms.

mark a long line approximately parallel to the optic axis and a short line indicating, in the case shown, the approximate direction and the sense of cut of a BT-plate. It is to be noted, here, that neither handedness nor electrical sense need be individually determined or considered, as such, for the sense of cut is directly obtained.

The size of an etch-figure depends upon the thickness of the section being examined, as explained in Section 5.43. For the etch-figures here presented the size of the figure relative to the thickness of the section, may be estimated

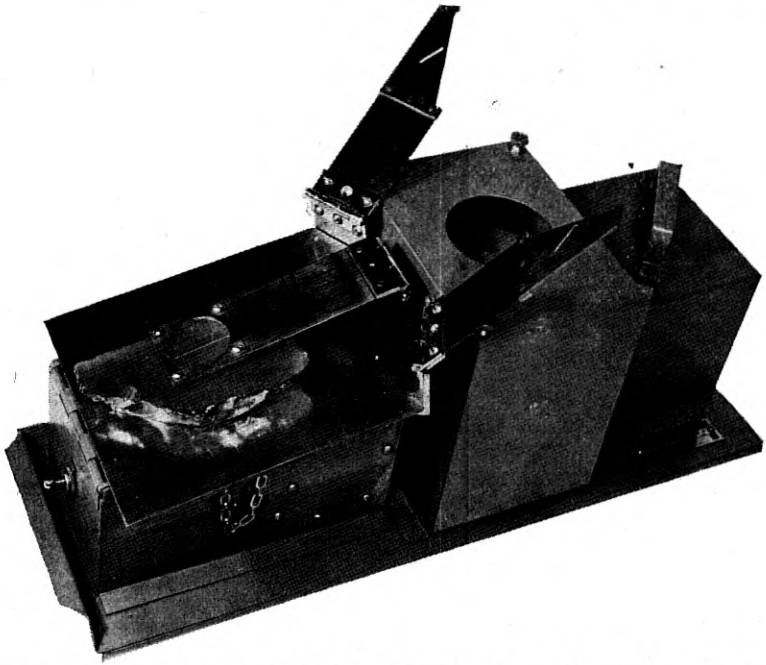

Fig. 5.18—The BASIC PIN-HOLE ORIASCOPE with matching and marking arms for use on X-cut sections. Twinning, and the direction of the Z axis, and the direction and sense of cut may be determined and marked on the section.

from a knowledge of the ratio, N, of the total diameter of the view to the thickness of the section giving that view. For Fig. 5.17A and B, $N = 1.3$; for Fig. 5.19A and B, $N = 2.7$; for Fig. 5.20, $N = 1.7$; for Fig. 5.21, $N = 2.5$.

The pin-hole oriascope may be used in a variety of other ways for examining any crystal cut with at least one etched surface. When used with sections as described above the bottom flat surface may be very small, just large enough to cover the pin-hole. However, this restricts the inspection

to an area directly over the bottom surface. This restriction may be eliminated, and no flat bottom surface need be used at all, if the bottom surface

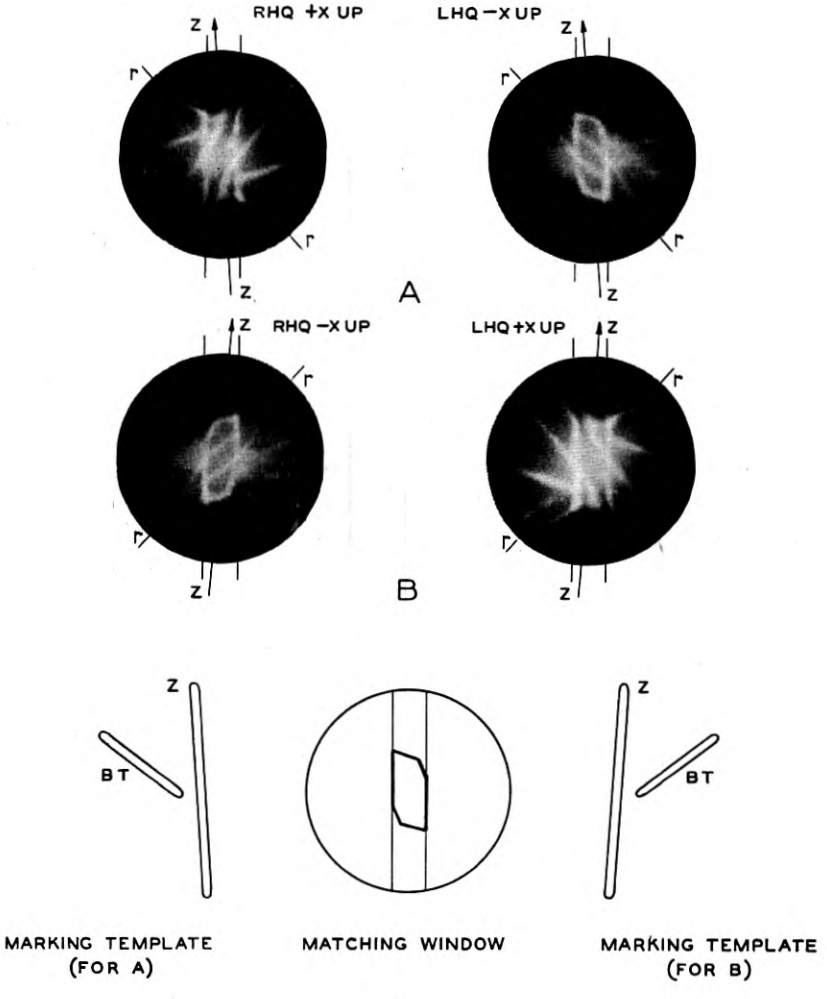

Fig. 5.19—Etch-figures obtained with the pin-hole oriascope in X-cut sections. After an etch-figure is aligned with the rulings on the matching window, as in C, the section is marked thru a marking template (in this case the one on the left) with the direction of the Z axis and the direction of cut of the desired plate (in this case the BT).

of the section be immersed in a transparent dish of immersion fluid (whose refractive index matches that of quartz) placed over the pin-hole. Here the

size of the etch-figure depends on the whole distance from the pin-hole to the etched top-surface, and hence, may be made as large as desired, by raising the section and fluid level. Very thin sections, slabs or plates may be examined similarly, with the bottom surface contacting the immersion fluid, or the plates may be wet with immersion fluid and placed on thick glass plates and placed over the pin-hole. In either case the top etched-surface must be kept dry. By this means the twinoriascope examinations described in Section 5.52 might be performed on the pin-hole oriascope, (a disadvantage being the necessity of using an immersion fluid).

Usually etch-figures are obtained from flat etched surfaces whose orientation is known within 5°. However, if the surface be 10° to 20° off-orientation the etch-figure will be plainly distorted. If now the section be viewed at an angle to the normal position, or if the section be tilted in the fluid-

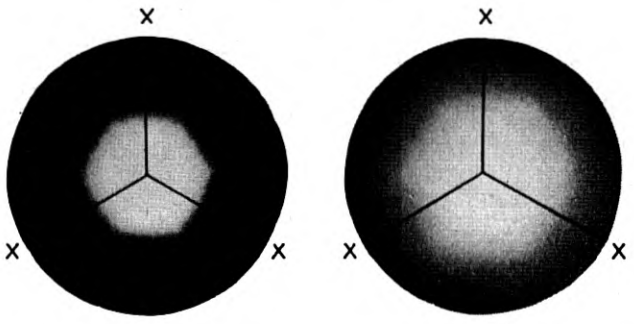

Fig. 5.20—CLEAVAGE-FIGURES may also be observed on the pin-hole oriascope in ground but unetched specimens, in this case a Z-cut section. Here the direction of the X axes but not their sense (nor handedness, nor twinning) may be determined.

bath method described above, the undistorted figure may be observed. The direction and amount of misorientation of the surface may be thus estimated. By provision of suitable mounts and scales the misorientation could be measured to 5°.

It might be added that in some cases unetched, ground (or sawn) quartz surfaces give "cleavage-figures." Thus with Z-cut sections which have been ground, but not etched, there may be observed on the pin-hole oriascope cleavage-figures like those shown in Fig. 5.20. The difference between the two views is mainly a difference in focusing and in photographic reproduction. The cleavage-figure indicates that there are preferential cleavage planes in quartz, which are parallel to the X-axes, and correspond approximately to the natural cap faces. Further, there is no indicated difference between the major and minor planes. Thus, the cleavage-figure is

six-fold and may not be used to determine electrical sense or twinning. It may, however, be used to determine approximately the orientation of the X-axes. Cleavage-figures are seldom strong, but appear to be best with coarse grinding.[18]

5.6 The Process of Etching Quartz

Few factors related to the chemical process of etching quartz have been extensively studied. Much of the information here presented is taken from preliminary reports of L. Egerton of the Laboratories, who has undertaken an investigation of the etching process. Though the information mainly regards hydrofluoric acid etching, some data is given on etching with hydrofluoric gas, and bifluoride mixtures.

The reaction of quartz, which is silicon dioxide (SiO_2), with hydrofluoric acid (HF) is given by the following equations:

$$SiO_2 + 6HF \rightleftarrows SiF_4 + 2H_2O + (2HF) \rightleftarrows H_2SiF_6 + 2H_2O.$$

Since the hydrofluoric acid is a solution of HF gas in water, the reaction of the acid with quartz results in a reduction of the concentration of HF. At the same time there is produced silicon tetrafluoride (SiF_4) which reacts with more HF to give fluosilicic acid (H_2SiF_6) in solution. It is common practice to start with about 50% HF acid and to continue etching until the HF concentration is down to 20 or 25%, at which time there should also be a 30% to 35% concentration of H_2SiF_6, if all the depletion of HF were due to reaction with the quartz. Actually much smaller concentrations of H_2SiF_6 are found, and this discrepancy is mainly due to the large continuous loss of HF from the solution by gassing. Further, the etching power of this used acid is not the same as would be obtained with a solution of 20%–25% HF alone in water. However, this difference is hardly noticeable except with weak etches.

Through the useful life of the acid, starting with 50% HF and depleting to about 20% HF, practically identical etch-figures may be obtained by properly adjusting the etching time. Means of testing the etching power of the acid to determine the proper etching time are complicated by the production of H_2SiF_6 in the solution, and by the irregular loss of HF by gassing. Further, the power of the acid to produce useable etch-figures is not the same as its power to remove quartz, or to etch glass, or as its concentration of HF or H_2SiF_6. For these reasons any indirect method of measuring etching-power must be correlated empirically with the etching-time required to give the desired etch-figures.

An indirect method of testing the etching-power, developed by Dr. W. Hoff of Western Electric, Hawthorne, involves the etching of sand blasted

[18] Scrubbing the surface with soap, water, and brush sometimes improves the figure.

microscope slides for a standard length of time. The lead-glass slides become coated with a white lead-fluoride deposit to a depth dependent mainly upon the HF content of the acid. The optical density of this deposit is measured with a specially adapted photometer. The photometer readings are correlated with required etching-times to give the desired etch-figures; a different etching-time being required for different kinds of sections, slabs, etc. Use of this means of controlling the etching time has greatly improved the regularity with which good etch-figures are produced in the shop.

Commercial hydrofluoric acid from a number of different suppliers has been analyzed for purity, and tested for the development of etch-figures. It appears that when such acids are brought to the same concentration (by addition of water if necessary) there is no difference in their effectiveness, nor are they inferior to pure reagent acid. Commonly the acid is supplied as 48% solutions in lead or hard rubber drums, or as 60% in steel drums (usually the concentration is a few per cent higher than labeled). The difference in packaging is of no importance in the results obtained, provided the concentration is properly reduced.

There are two important factors regarding the starting concentration of hydrofluoric acid baths. In the first place, acids stronger than 50%, though reacting vigorously with the quartz (and removing material rapidly), do not give good etch-figures. Secondly, strong acids not contained in sealed containers lose strength very rapidly by gassing of the HF gas. Hence unused fresh acid should be kept well stoppered. Before use the acid should be diluted to a concentration of 45% to 50%. This may be accomplished by adding about $\frac{1}{3}$ volume of water to one volume of 60% acid, or $\frac{1}{6}$ volume of water to one volume of 55% acid.

Concentrated hydrofluoric acid loses HF by gassing more rapidly than it loses water by evaporation. This preferential loss of HF continues until the HF concentration is reduced to 35% or less,[19] and is not completely overcome by covering the bath without sealing. In fact, in practice, it appears that about as much HF is lost by gassing as is used in etching the quartz. Thus the bath should be kept as tightly covered as is practicable.

Whereas, in the past only lead and hard rubber have been used for fabrication of acid baths and racks, it appears that for concentrations not greater than 50% HF, copper, nickel, and brass may be used as well (steel is inferior at low concentrations). Lead-tin solders may not be used, but silver solder is satisfactory. Thus shop acid equipment may be easily fabricated out of common fabricating materials.[20]

[19] At room temperatures there appears to be a constant-concentration mixture at some concentration below the 35% concentration of the constant boiling mixture, the exact value depending upon the temperature of the solution and the ambient humidity.

[20] Polystyrene is a good material for use in fabrication of vessels for handling HF and its reaction products in the laboratory.

While agitation of the acid bath during etching does speed up the removal of quartz from the surface, it does not appear to speed up the development of the etch-figures here considered. However, moderate agitation does improve the uniformity of etch from one crystal to another, and even over the surface of single large surfaces (especially when such surfaces are close together). Uniformity of etch is important in examining for twinning. The surfaces to be etched should never be placed in contact with each other, or with other surfaces, so that the acid cannot flow between them (the separation should be at least $\frac{1}{32}$ of an inch).

The effect of temperature on the etching process appears to be small for the range of room temperatures normally encountered in practice.

A word of caution chould be added regarding the handling of hydrofluoric acid and other fluorine etching materials. The dangers are of two kinds. First, fluorine poisoning may result from contact with any fluorine compounds, the effects of which may be cumulative. Special care should be taken to prevent inhalation of vapors from all etching baths containing fluorine. Some persons are especially sensitive to fluorine poisoning. Secondly, hydrofluoric acid baths, or any baths containing free HF, may produce acid burns. Commonly such burns are attended by fluorine poisoning. For these reasons etching with all fluorine compounds is preferably carried out in ventilated hoods (with strong air suction through the door), with continually running water for washing, and with rubber gloves, tongs, racks, etc. for handling the quartz.

Etching compounds other than hydrofluoric acid have been widely used in etching glass, as is evidenced by the variety of formulae presented in the "Chemical Formulary."[21] Solutions of ammonium bifluoride (NH_4HF_2), with additions of various amounts of free hydrogen fluoride, sodium bifluoride, sugar, and other materials have long been used on glass. One of the possible advantages of such formulae for etching quartz is the elimination of the dangers of acid burns and strong fumes that may be obtained with hydrofluoric acid (care must still be maintained to prevent fluorine poisoning). A number of these formulae have been made up and tested on quartz. The preliminary conclusions are as follows.

The etch-figures that may be developed by the bifluoride compounds on Z and X-cut sections of quartz are not the same as those developed by hydrofluoric acid. The results approach each other, however, for excessively long etching in both cases. To obtain *usable* etch-figures on X-cut sections with the bifluoride requires considerably longer etching time than with hydrofluoric acid, or an elevation of the bath temperature to about 45°C. The addition of hydrofluoric acid to the bifluoride formulae speeds up the development, but partly negates the safety advantage of the bifluoride bath.

[21] Published by the Chemical Publishing Co., Brooklyn, N. Y.

The figures produced on Z-cut surfaces are small and complex, (hardly usable) unless a considerable amount of free HF acid is added. Etch-figures here considered are those produced on the pin-hole instrument, and are usable only if they have such character as will permit of their use in determining quartz axes. Fig. 5.21 shows the type of usable etch-figure obtained on X-cut sections with an ammonium bifluoride and sugar solution (the sugar is here effective mainly in preventing creepage of the solution). It might

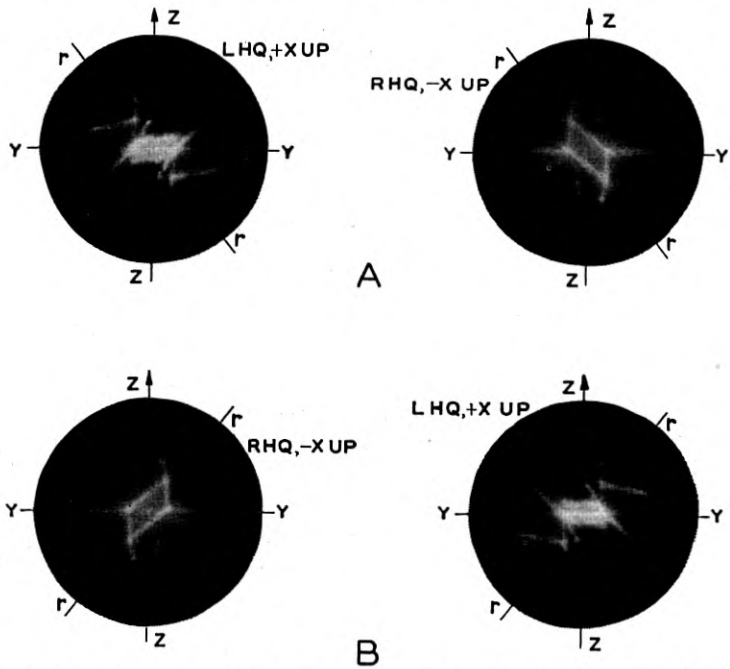

Fig. 5.21—Etch-figures obtained on the pin-hole oriascope with X-cut sections which have been strongly etched in bifluoride mixtures, or excessively etched in hydrofluoric acid. These etch-figures differ from those of Fig. 5.19 (for a moderate etch in hydrofluoric acid) but are obviously usable.

be noted that a similar figure is obtained with hydrogen fluoride gas, and with *excessively* long etching (several hours) in hydrofluoric acid.

When the bifluorides are used only to develop reflection contrast in the detection of twinning, their effectiveness appears to be about the same as hydrofluoric acid, under equivalent process conditions. The etching power of the bifluorides may be maintained nearly constant over a long period of use by maintaining an excess of the salt in solution, a distinct advantage over the acid. The metals copper, nickel, brass and stainless steel may be used in fabricating tanks and racks, lead and steel are inferior.

Finished quartz surfaces are sometimes etched to remove surface debris

(fragments of quartz loosened by grinding, and grinding refuse embedded in microscopic surface irregularities), and to remove predetermined small amounts of the surface for frequency adjustment. It is common for these purposes to use weaker etching solutions, since very small amounts of quartz are to be removed. With hydrofluoric acid, weak solutions (less than 20% HF) have an advantage in that their concentrations are little reduced by exposure to the air. In fact with very weak solution the concentration may increase slightly by exposure, and thus partly compensate for the HF lost by reaction. Weak ammonium bifluoride solutions may also be used, provided no deposit forming material is added.

5.7 The Effect of Twinning in the Finished Plate

While it is commonly considered that electrical and optical twinning are not allowable in a finished oscillator plate, it cannot be unconditionally stated that small amounts of twinning will too seriously affect the properties of all types of oscillator plates. The allowance of even small amounts of twinning in the finished plate would save quartz and simplify the processing procedures. Hence, consideration must be given to the factors which would affect the utilization of twinned material, and the effect of twinning on the operating characteristics of the finished plate. Consideration will first be given to the nature and distribution of electrical and optical twins[22] in the raw quartz.

The analysis of twinning in raw quartz has been carried out by the examination of numerous, etched Z-cut surfaces. By the method to be described it is possible to detect the handedness, and the axial orientation and sense, of each homogeneous portion, twin, appearing at the etched surface of a twinned specimen. Both electrical and optical twins may be analyzed by this method. It might be added that electrical twinning boundaries and orientation are only detectable at an etched surface, and that while interior optical twinning may be detected by polarized light, its exact analysis is only possible at an etched surface.

Figure 5.22 E shows the optical arrangement used for examining twinning in etched Z-cut sections. The sections (prepared with a fine grind and weak etch) were mounted on a turntable, illuminated from an elevation of about 30° to the horizontal etched surface by a spot lamp, and viewed (or photographed) from vertically above the section according to principles of Section 5.41). With the section properly aligned on the table (with the predetermined electric axes parallel to the table-lines joining diametrically opposite fiducial marks), the table was successively turned into positions about 12° to the right or left of the plane of illumination and reflection (as indicated by the R and L marks and the index pointer). Four of these positions of

[22] See footnote 4.

illumination of a given section are sufficient to determine the nature of the four possible twins in the section. The four corresponding photographic views of the section have been arranged in a special manner to simplify their explanation. This arrangement, as shown in Fig. 5.22A, B, C and D, is equivalent to what would be observed if one looked down on a single, stationary section, and illuminated the section from the four different direc-

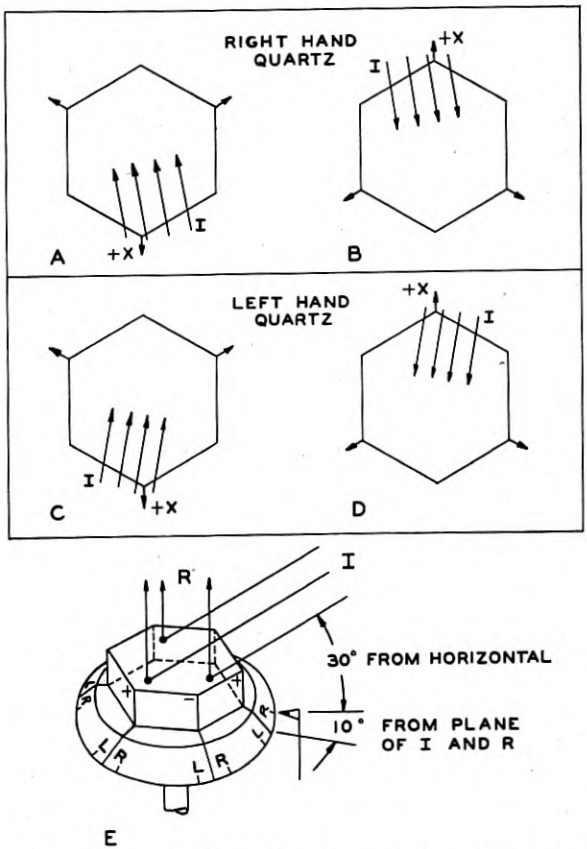

Fig. 5.22—Reflection patterns of the twinned, Z-cut sections shown in Figs. 5.23, 5.24, 5.25 and 5.26 were obtained by the means shown in E. A, B, C, and D are a key to the four equivalent directions of illumination of a single stationary section.

tions shown in the figure. For each direction of illumination there is a corresponding view, the outline of the section (and any cracks, chips or other flaws) being identically positioned in each view. However, when the four types of twins are present in a given section, each view will show a different region, or regions, of brightness. For each view, the interpretation of handedness and electrical sense of the *bright portion* of the view is according

to the labeling of this particular view, only. Thus if a section is entirely right quartz and of the electrical sense shown at A the whole surface of the section will appear bright in view A and dark in all other views. If a section is all right quartz, but partly of the electrical sense shown in A and partly that shown in B, then part of the surface will appear bright in A and the other part will be bright in B (the whole surface will be dark in C and D).

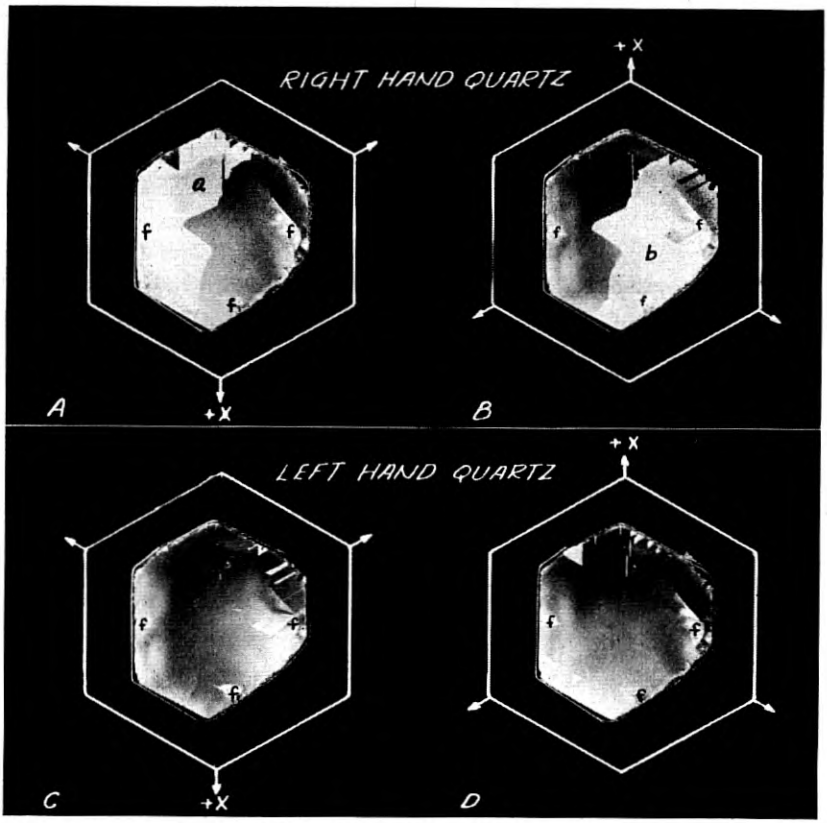

Fig. 5.23—The four possible conditions of handedness and electrical sense in a single section are shown here. In each view the handedness and sense is for only the *bright* portion of that view. The *a* and *b* regions are seen to be both of right quartz but of opposite electrical sense, hence electrical twins. (Flaws indicated by *f* are to be disregarded).

A section containing all four possible twins would exhibit bright regions in each view, and a different bright region in each view. All bright regions would fit together to make a complete map of the surface. Only the bright portion of each view has the handedness and electrical sense indicated for that view.

Figure 5.23 shows a Z-cut section containing twins of the four possible

conditions of electrical sense and handedness. The two large, bright regions *a* and *b* (appearing in views A and B respectively) are both right quartz but of opposite electrical-sense. Hence the surface is mainly of electrically twinned right quartz. The small dark regions within the borders of *a* (view A) are bright in view D. Hence these small, triangular and line regions are left quartz and of opposite electrical sense to the large region *a* containing them. They are then optical twins of the large *a* region. Similarly the dark regions of *b* (view B) are found from view C to be optical twins of the *b* region. (Flaws labeled *f* are cracks, chips, etc.) If the whole section were cut up to make AT plates, for example, and at the proper angular sense according to the *a* portion of the section, then those plates coming from the *b* region would be of wrong angular sense. Those crossing a boundary between the *a* and *b* regions would be of both senses, i.e., electrically twinned. Those few plates which contained some left quartz would be optically twinned. To make the most economical use of this section it should be separated, by cutting along a line approximating the *a* to *b* boundary, so that each half of the section may be cut at the correct sense of orientation. Even when so cut, some of the plates will contain optical twinning and remnants of electrical twinning. This section is typical of much of the raw quartz that must be used for manufacturing piezoelectric plates.

Figure 5.24 shows a section which is mainly of left quartz as exhibited by the large bright *c* and *d* regions of views C and D. The large *c* region is optically-twinned to a small extent by the line regions *b* of view B. One of the *d* regions is badly optically twinned by the small striated *a* regions, as seen in A. Such a section would be very uneconomical to process, since separating the larger electrical twins is not feasible. If processed at all, it should probably be entirely cut according to the handedness and sense of the large *c* portion, the wrong-sensed regions and twinning being cut away at a later stage (after inspection of the slabs in the twinoriascope, for example). It might be noted that only the optical twinning could have been observed in the initial polarized-light, raw quartz inspection, where such a stone would be passed as moderately good.

Fig. 5.25 shows an *unusual* section that is mainly composed of left quartz, regions *c* and *d*. The right quartz regions shown in view B are of both opposite-handedness and electrical-sense to the *c* region inclosing or bordering them. This is the common and expected conditions. The unusual condition is exhibited by the regions *c* and *a*, where twins of opposite-handedness but *same* electrical-sense have a common boundary. Since this boundary could be detected by optical means, the *a* and *c* regions might be described as optical-twins, of an "uncommon variety". However, by convention *optical twinning* has long been used to denote twinning exhibiting both opposite-handedness and opposite-electrical-sense (crystallographically,

Brazil twinning). Further, twinning exhibiting both opposite-handedness and same-electrical-sense, combines the crystallographic twinning laws of Brazil twinning and Dauphiné (electrical) twinning. Hence this uncommon variety of twinning may preferably be called combined electrical and optical twinning, or just COMBINED TWINNING. Thus, the boundary

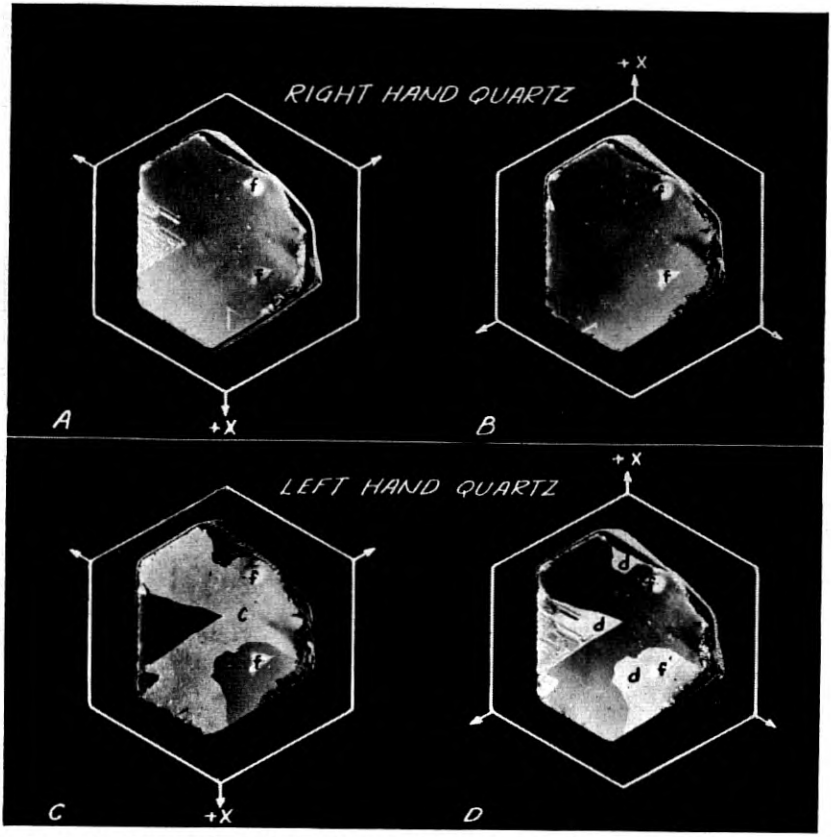

Fig. 5.24—Regions *d* are electrical twins of the region *c*. The striated regions *a* are of opposite handedness and electrical sense to the *d* region enclosing them, hence optical twins of *d*. The *b* regions are small optical twins of *c*, and *f* are flaws.

between the *a* and *c* twins separates combined twins. Note also that the *a* twin bounds the *b* twin and the *b* twin bounds the *c* twin. Thus, *a* and *b* are true electrical twins, and *b* and *c* are true optical twins.[23]

[23] It is possible that growth conditions are such that combined twinning cannot occur by itself, without the presence of true optical twinning and true electrical twinning. That is, a region of given handedness and sense can not be entirely bordered by a region of opposite-handedness and same-sense.

Figure 5.26 shows an unusual section which is mainly composed of left quartz, of the electrical sense shown in D, region *d*. The region *c* is an electrical twin of *d*. The region *f* is a flaw in the quartz and is to be disregarded. The region *a* is an optical twin of *d*, and is *uncommonly* large for an optical twin (note: region *a* contains within it, two small optical twins). Since

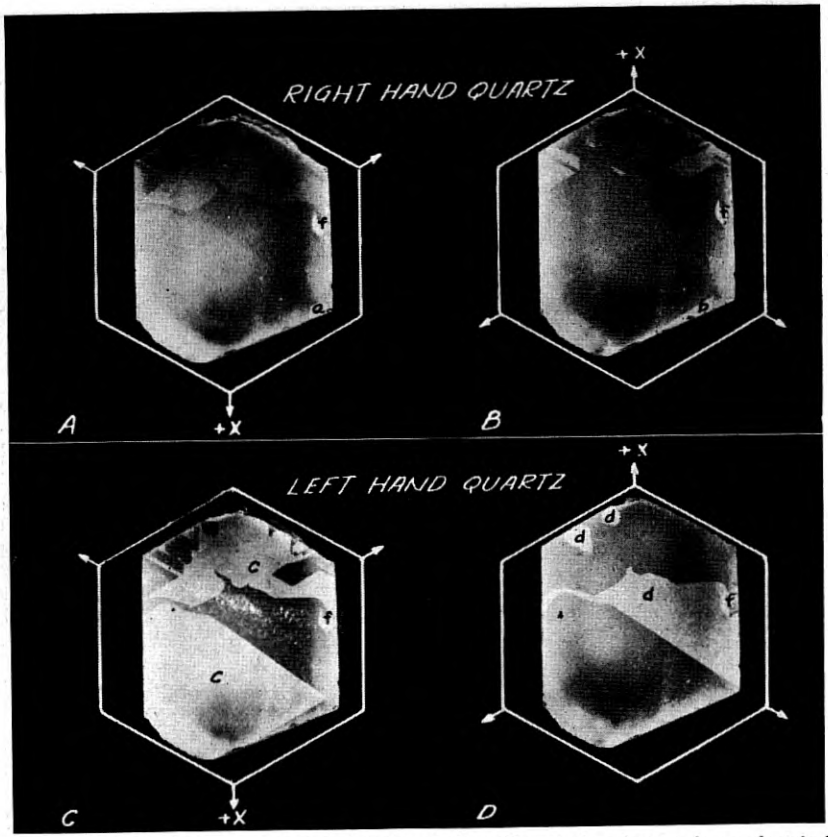

Fig. 5.25—Regions *c* are electrical twins of the adjacent *d* regions, *a* is an electrical twin of *b*, and *a* is also an optical twin of *d*. An uncommon condition of twinning is presented by the adjacent *a* and *c* regions which are of opposite handedness but the *same* electrical sense, thus exhibiting COMBINED-TWINNING.

optical twins are usually very small (except for the one major surrounding twin), it is seldom possible to cut them apart and use each twin individually.

Figures 5.1, 5.2 and 5.3 were obtained by the means above described, and all sections shown in these figures (except Fig. 5.2A and C) actually exhibited both electrical and optical twinning. Thus Fig. 5.3D was obtained from Fig. 5.24A, and Fig. 5.2F from Fig. 5.24C, etc., by trimming the latter

named figures to give the sections simulated natural faces. Figures 5.2 and 5.3 are of particular use in learning to distinguish between electrical and optical twinning when examining etched surfaces by reflection. Note that electrical twins are usually large and separated by irregular boundaries, Fig. 5.2. Optical twins are usually separated by straight-line boundaries

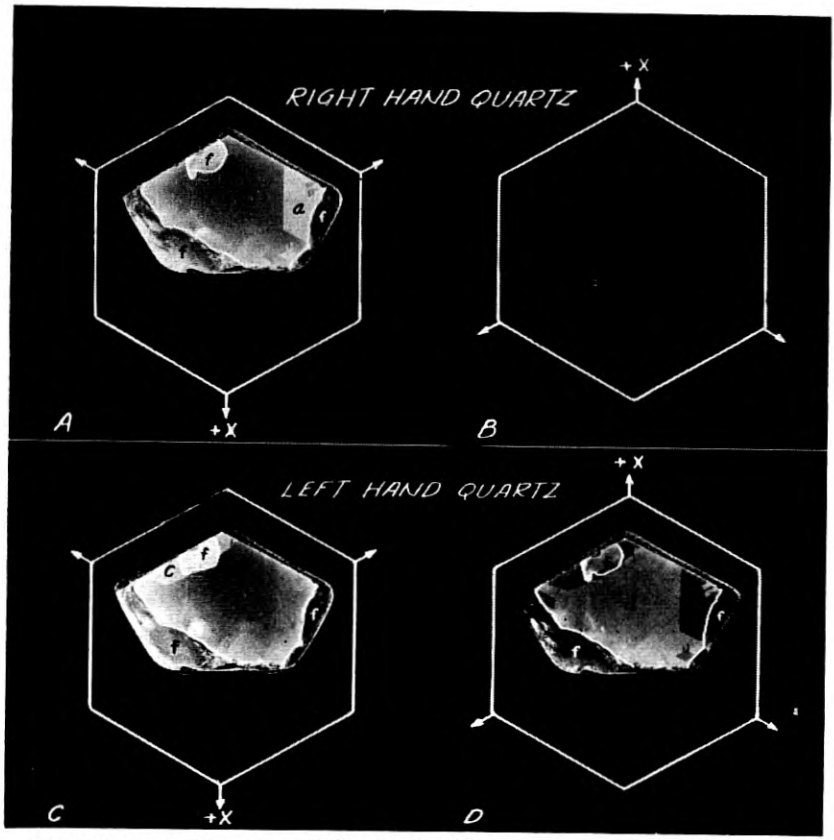

Fig. 5.26—Since this section exhibited no bright regions (except flaws *f*) in view B (i.e. no right quartz of electrical sense B) it was not reproduced in view B. The *c* region is an electrical twin of the adjacent *d* region, while *a* is an optical twin of *d*. It is *uncommon* for a minor optical twin to be as large as *a*.

parallel to natural faces, thus forming triangular, parallelogram, and straight line insets, Fig. 5.3. Optical twins (except for the one major, surrounding twin) are usually very small and often interlayered (with the major twin). Large interlayered regions are entirely unusable and hence are cut away at the earliest possible stage to save the labor of processing worthless material.

Small optical twins and small electrical twins (or remnants of electrical

twins left after cutting electrical twins apart) may be isolated or removed in an intermediate or late stage of processing, where they are detected by the etch technique. Commonly the final rejection of material twinned in either way is delayed until after the final blanks are cut out. These may be etched and examined by reflection, one at a time under a spot lamp, and those showing twinning (and other imperfections) sorted out and rejected.

Another possible method of rejecting twinning which is of sufficient amount to be harmful is by making electrical tests on the finished (or semi-finished) plates, at which time those plates failing to meet the electrical tests for any reason (including twinning), are rejected. While this method of rejection does not assure that twinning will be entirely absent from the accepted plates, neither does any other method assure complete absence of twinning. Further, except for imperfections which may affect the useful life of the plate, acceptance of finished oscillator plates is not illogically based

TABLE I.—*Constants for Plates of Correct and Incorrect Sense of Cut*

Cut, Angle	Frequency Constant (fxd. in Kc. mm.)	Temperature Coefficient (parts/10^6/C.°)
AT +35° (−35°)	1670 (2400)	0 (+30)
CT +38° (−38°)	3080 (2100)	0 (−30)
BT −49° (+49°)	2560 (1880)	0 (−55)
DT −52° (+52°)	2060 (2850)	0 (+45)

upon their meeting the desired electrical operating characteristics, i.e., frequency, temperature-coefficient, activity and internal damping (all determinable by electrical tests).[24] It does not appear that twinning will affect the useful life of the plate. Its effect upon the electrical operating characteristics of the plate depend upon many factors.

An important factor regarding twinning in the finished plate is that optical twinning introduces a less important variation in the physical properties of the plate than does electrical twinning. Thus, in the case of optical twinning alone, both portions of the plate are of the same sense of cut, though still being of opposite electrical sense. This may be understood from an examination of Fig. 5.4, the second and third views taken together represent optical twinning. In the case of electrical twinning the two portions of the

[24] With filter plates additional operating characteristics must be met. The ratio of capacities (see Chap. I, Appendix A.3) is greatly affected by the opposed electrical sense of twinning.

plate are of both opposite sense of cut and opposite electrical sense, as may be observed from the third and fourth views of Fig. 5.4. The effect of this difference in sense of cut for the two types of twinning is brought out by Table I, which gives the approximate frequency constants and temperature coefficients for the common cuts of oscillator plates, together with those for the analogous, oppositely (and hence wrong) sensed cuts.

In the case of a CT plate, for example, both portions of an *optically twinned* plate (cut at +38°) will be of the same +38° orientation. The plate is elastically the same throughout and hence should exhibit the frequency and low temperature-coefficient desired. However, the opposed electrical senses of the two portions will cause a reduction in the electrical activity. The amount of this reduction will depend upon the relative size of the two portions and upon their placement relative to the vibration nodes of the plate.

On the other hand, when a CT plate is *electrically twinned* one portion of the plate will be of the correct +38° orientation while the other portion is of the incorrect −38° orientation. The two portions of the plate have widely different elastic properties, as is exhibited in the table by the different frequency constants and their respective temperature-coefficients. Resulting from this difference alone, the plate will exhibit operating characteristics (if operable at all) intermediate between the two listed in the table (usually near one of these two), and its activity will be reduced. The activity will also be reduced by the opposite electrical senses in the two portions. The degree to which the frequency, temperature-coefficient, and activity are affected, again depends upon the relative sizes of the two portions of the plate and their placement relative to the "nodes" of the plate.

Thus, for equivalent proportions and placement of twinning, electrical twinning will cause a much greater change in the operating characteristics of the plate than will optical twinning.[25]

A note may be inserted regarding the electrical testing of plates, some of which may be twinned while others may be untwinned but of incorrect sense of cut. As seen from Table I, untwinned plates of the correct sense of cut are easily distinguished from those of the incorrect sense of cut by their frequency. This distinction between sense of cut holds as well for plates containing very little twinning. The presence of appreciable twinning in the plate is easily distinguished by the activity of the plate. While ordinarily a plate would be electrically tested in the mode of vibration it is intended to be operated in, it is sometimes of advantage to test it in a different mode.

[25] In the case of the *uncommon* "combined-twinning" the two portions of the plate are of opposite sense of cut but of the same electrical sense. The effect on the operating characteristics will be like that for electrical twinning, except that the activity may not be as greatly reduced.

Thus the high-frequency mode plates (AT and BT) might be tested in their low frequency modes (corresponding roughly to the CT and DT modes, respectively). A further discussion of this matter will be found in a later chapter by I. E. Fair.

5.8 Conclusions

In the processing of quartz, consideration must be given to the nature of twinning and to its characteristic distribution in the raw stone. There are only two common types of twinning that need be considered, namely *electrical twinning* and (true) *optical twinning* ("combined-twinning" and other types are a rarity). Due to the characteristically large size (and the nature) of *electrical twins*, a stone must be examined for electrical twinning (by the etch technique) at an early stage of processing so that the electrical twins may be observed and cut apart before the angular cuts (AT, BT, CT, DT, etc. slabs, bars, or wafers) are made. Otherwise, some of the large electrical twins will be entirely cut up with the incorrect angular sense, and hence wasted.

On the other hand *optical twins* are characteristically small and interlayered, or small and scattered. The interlayered regions are entirely unusable. Hence processing labor will be saved by inspection of the raw stones (by the polarized light means of Chapter IV), and of the first sections at least (by the etch technique) for large regions of interlayered optical twinning.

Scattered optical twins and small electrical twins, or remnants of electrical twins which have been cut apart, may be cut away in an intermediate processing stage, or in a later stage plates containing such twinning may be separated out. In either case the etch technique may be used to detect the twinning.

An alternative method of eliminating small electrical twins (or remanents thereof) and of small optical twins (most of which are characteristically very small) is by electrical tests on the finished plate. This method has merit in that if the twins are sufficiently small, and not disadvantageously placed in plate, they may not harmfully effect the desired operating characteristics of the plates. The degree of the effect depends not only upon the size of the twin and its location in the plate, but upon whether the twinning is electrical or optical; *optical twinning* being considerably *less harmful* than electrical twinning. The effect of the twinning further depends upon the type of plate being considered, i.e. its size and mode of operation, and use. It is probable that twinning is more tolerable in low-frequency mode oscillators (CT and DT) than in the high frequency modes (AT and BT), and of course more tolerable in plates of low requirements on the operating characteristics (activity, frequency and temperature-coefficient). Twinning is

probably least tolerable in filter plates, which have to meet very special requirements.[26] Detailed experimental studies of allowable amounts of twinning are of little value since to use the results in a manufacturing process would require a careful inspection of each plate and a difficult classification into groups depending upon the variety, amount, and placement of the twinning. Acceptance or rejection of finished plates on the basis of their final electrical operating characteristics appears to be the only practical means of separating usably twinned plates from unusably twinned plates. This method of selection does not determine whether the rejected plates contain twinning or other imperfections (or are misoriented or misdimensioned) and is therefore of little use in analyzing the processing methods to determine best practices. This disadvantage may be eliminated by etching the *rejected* plates and examining them for twinning (and such other imperfections as show up best after etching).

The effects of crystal imperfections other than twinning were discussed in Chapter IV, Section 4.9.

[26] See footnote 24.

CHAPTER VI

Modes of Motion in Quartz Crystals, the Effects of Coupling and Methods of Design

By R. A. SYKES

6.1 Introduction

WITH the recent extended use of Quartz crystals in oscillators and electrical networks has come a need for a comprehensive view of the various types of crystal cuts. In addition there has been a need for illustration of some of the methods employed in choosing the proper cut for a given requirement, the manner in which quartz crystals vibrate and the basic principles governing the choice of a design to use certain cuts most advantageously. In particular one of the greatest problems associated with the recent large scale production of crystals for oscillator purposes has been that of obtaining crystals the activity and frequency of which would not vary to any large degree over a wide range in temperature.

It is the intention of this chapter to present a physical picture of the manner in which quartz crystals vibrate in their simplest forms and then to show what has been learned from these simple forms that will apply to the more complex combinations of motion. The motion of a bar or plate is determined almost wholly by its dimensions and the particular type of wave generated, or frequency applied, and very little upon the driving system if the coupling to the driving system is small. In the case of quartz the coupling between the electric and mechanical system is small and hence we may study the motion of rods and plates without always considering the effect of changes due to the method of excitation (i.e., piezo-electric). However the ease of exciting and measuring a particular mode does depend on the piezo-electric constant driving it. Basically only three types of motion will be considered; flexural, extensional and shear. These three types of motion or combinations of these can be considered to represent most of the cases with which we will concern ourselves. In addition, the frequency equations will be given for common types of motion and the effect of coupling between various modes of motion. Finally the general rules relating to the dimensioning of oscillator plates will be presented.

6.2 Types of Motion in Quartz Rods and Plates

6.21 *Flexural*

The motion associated with flexure will be discussed first because this is the type of motion that we see more commonly in nature. This motion is

the type which presents itself in the xylophone, the chime type door bell, and various other vibrating reeds or bars. Fig. 6.1 shows the general type of motion of a bar free to vibrate in flexure. The displacement takes place in the direction of W and the wave is propagated along the length. A flexure mode is one in which the center line does not change length. The type of motion associated with the first order, or fundamental, of a bar free to vibrate on both ends is shown in Fig. 6.1 with a dotted figure superim-

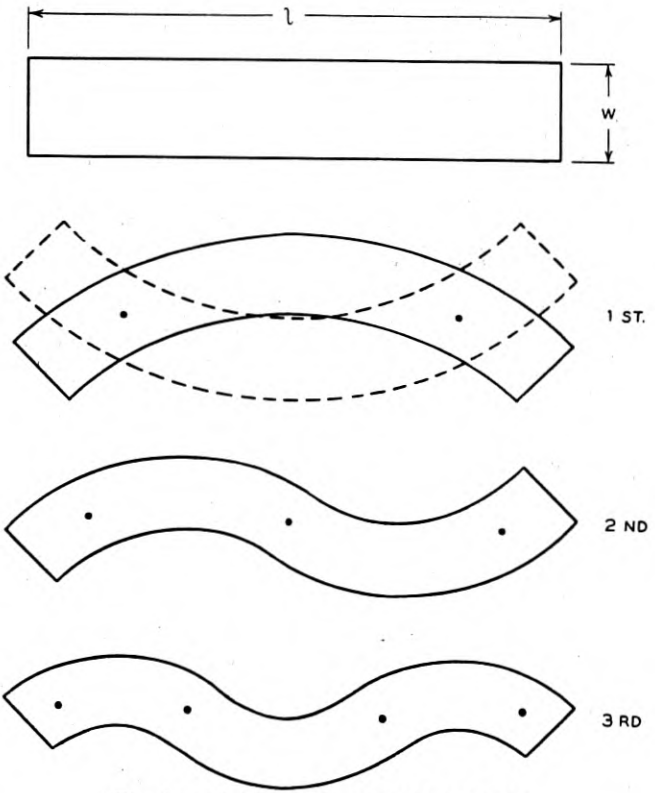

Fig. 6.1—Motion of a bar in free-free flexure.

posed to show the motion in the opposite phase. The straight bar then would be distorted first in one direction and then in the direction of the dotted figure. In the case of the second mode of vibration, it will be noticed that it consists essentially of two of the fundamental mode types joined end to end. This is not strictly the case, but serves to illustrate the motion. The dots shown at various points on the bar show positions of zero motion or nodes. In the case of the fundamental mode, there are two nodes and in the second and third there are three and four respectively. One point of

interest in flexure vibration as seen in Fig. 6.1 is that the ends of the bar will be vibrating in the same direction for odd order modes and the motion of the two ends will be in opposing directions for even order modes. The frequency of a bar vibrating in flexure may be easily computed for low orders when the width is small in comparison with the length. When the width is appreciable other factors must be considered as will be shown later. In general, the flexure frequency of a bar will be the lowest frequency of vibration.

In the case of a plate where we are concerned with flexural vibrations propagated along the length with motion in the direction of the thickness it

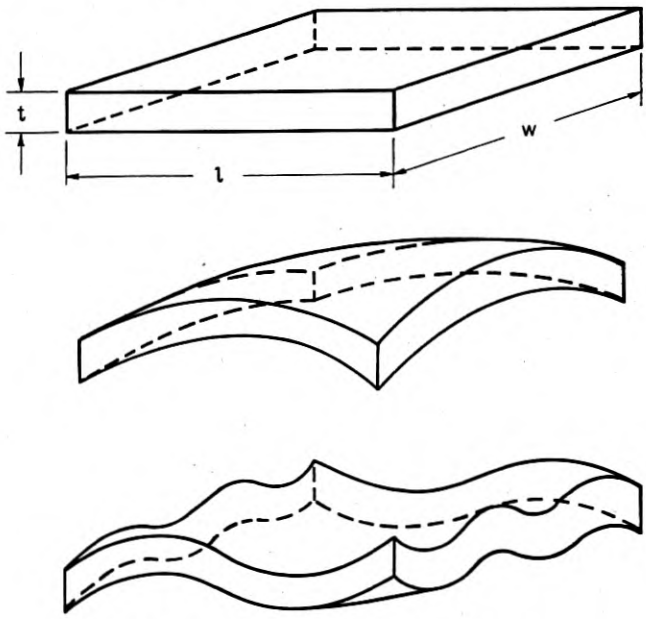

Fig. 6.2—Motion of a plate in free-free flexure.

is necessary to consider also the width. As noted in Fig. 6.1, our concern was only for a bar of small third dimension. When considering the case of a plate in flexure along its length and thickness, then the third dimension must also be considered for more complicated types of motion. In a manner somewhat similar to the vibration of a bar, we can consider a plate vibrating in its thickness-length plane. Since a plate also has width, we must also consider this dimension. The simplest type of motion would be that of a simple flexure which would bend the plate into the shape of an arch. If now, the third dimension is permitted to flex, the distortion of a plate shown in Fig. 6.2 could be illustrated by a flexure in the ℓ–t plane and in the

w–t plane. Considering the motion of the plate as a flexure vibration along the length vibrating in the thickness, then we may also have a distortion along the width and thickness corresponding to similar or higher types of flexure motion. The illustration at the bottom of the figure shows a plate vibrating in its second order flexure along the length and thickness and the fourth order flexure along the width and thickness. The effect of these higher orders in the w–t plane is to slightly modify the frequency of the ℓ–w mode.

A thorough treatment of this type of double flexure in plates will be given in Chapter VIII by H. J. McSkimin.

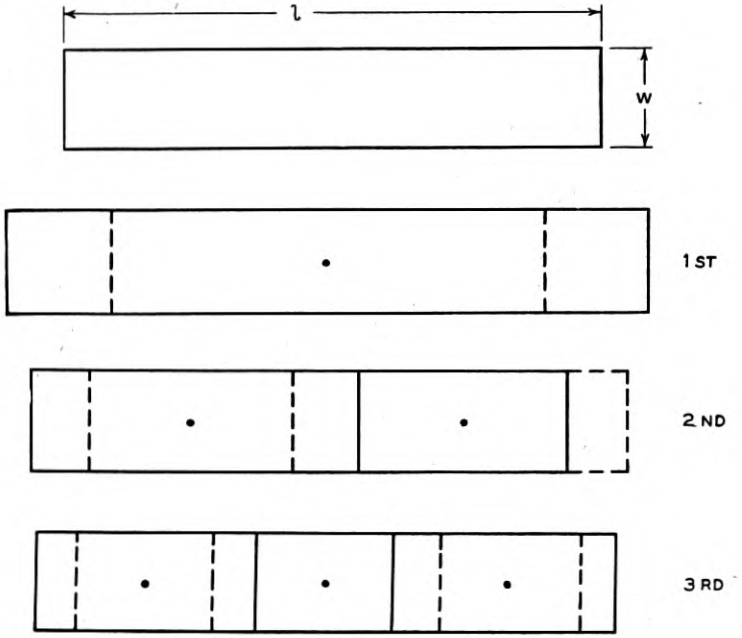

Fig. 6.3—Motion of a bar in free-free extension.

6.22 *Extensional*

The extensional or sometimes termed longitudinal motion of a bar free to vibrate is shown on Fig. 6.3. This motion is somewhat simpler than the flexure motion and consists simply of a displacement in the direction of the length of the bar of a wave propagated along the length. This means that the first mode of vibration will be simply an expansion and contraction of all points with respect to the center of the bar. This motion will be along the length. The displacements along the bar will then be in proportion to the sine of the angular distance from the center. The distortion of a free bar in its simplest mode is then illustrated in Fig. 6.3 labeled 1st. Since the

motion must be dynamically balanced, a node will appear at the center of the bar, and the bar will grow longer and shorter as shown by the solid and dotted lines. In the case of the second order of motion, as shown in Fig. 6.3, it consists essentially of two 1st order modes joined together at their ends and of opposite phase. That is to say, when one half of the bar is expanding, the other half is contracting. In the case of the 3rd mode, as can be seen from Fig. 6.3, the central element is contracting while the external elements are expanding. From this we may state generally, that for odd order types of motion, the extreme ends of the bar will be expanding or contracting in phase and for even order modes, the extreme ends will be expanding or contracting in opposite phase. Fig. 6.3 illustrates extensional motion in its simplest form. In a practical case an extension in one direction is accompanied by a contraction in one or both of the other two dimensions. This of course is due to elastic coupling and will be considered more in detail later. If we consider a rectangular plate it is not difficult to imagine that it would have three series of extensional modes of vibration due to the three principal dimensions.

6.23 *Shear*

The low frequency of face shear type of motion of a plate is somewhat more complicated than either the flexure or longitudinal and, as shown in Fig. 6.4, consists simply of an expansion and compression in opposite phase along the two diagonals of the plate. This motion is shown in Fig. 6.4 labeled $m = 1, n = 1$. The two phases are shown, one a solid curve and the other a dotted curve to illustrate the distortion with respect to the original plate. One peculiarity of shear motion in plates is that it may break up into motions similar to its fundamental along either the length or the width. For example, if we take the motion associated with $m = 1$, $n = 1$, and superimpose two of these in opposite phase on the same plate, we would get the type of motion illustrated by $m = 2, n = 1$. In a similar manner, the motion may reverse its phase any number of times along either the length or the width. One particular case is shown for $m = 6, n = 3$. As can be seen from the case of $m = 1, n = 1$, the distortion is not that of a parallelogram as it is in the static case because here we are concerned only with the dynamic case. While the equation of motion of a free plate vibrating in shear has not been completely solved, a microscopic analysis indicates that the actual motion of the plate edges appear to be somewhat as shown for the case $m = 1, n = 1$ when driven in this mode.

The shear mode of motion in the case of a thin plate is somewhat different for the high frequency case than for the low frequency case. In the case of high frequency shear modes of motion in thin plates, the motion of a particle is at right angles to the direction of propagation which in this case would be

the thickness. The simplest type of motion for high frequency shear is shown in Fig. 6.5 where the top of the plate is displaced in the direction along ℓ with respect to the bottom of the plate. This would then be termed the length-thickness shear. When viewed from the edge of the plate, the motion is very similar to that shown in Fig. 6.4 for the case of $m = 1, n = 1$. In a manner similar to the previous case of shear the front edge of the plate may be divided into segments along ℓ and along t. For example, we may get

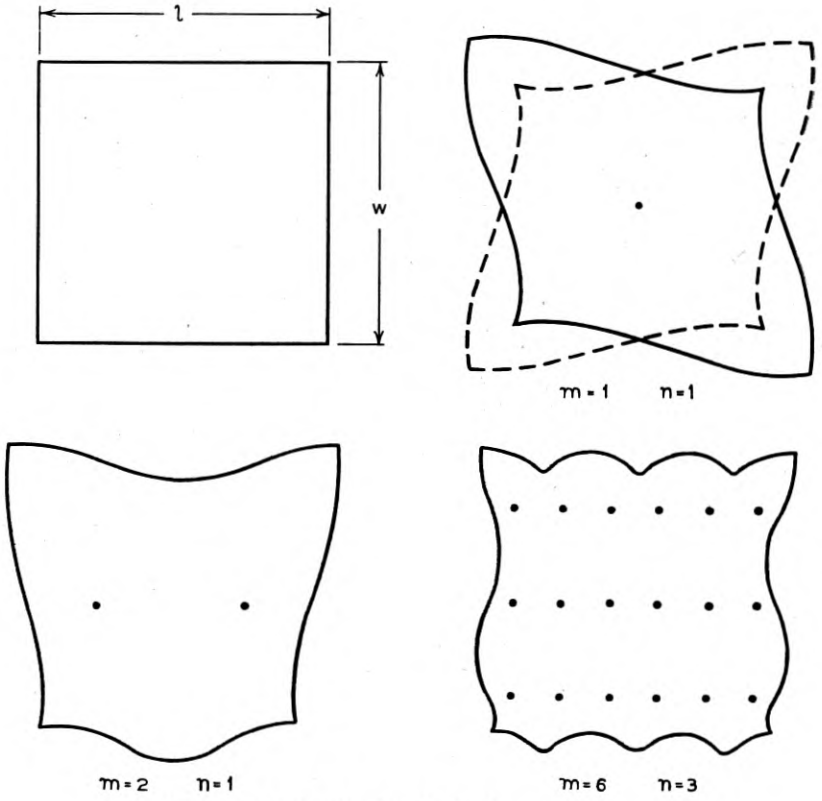

Fig. 6.4—Motion of a plate in low frequency shear.

a double shear along ℓ with a single shear along t. This case is illustrated in Fig. 6.5 for $m = 1, n = 2$ and $p = 1$. In general, m and n may assume any integral value. As in the case of flexure we must also consider the third dimension. The motion associated with the third dimension may be represented by simple reversals of phase as before. For example, in Fig. 6.5 the case for $m = 1, n = 1, p = 2$ is shown which simply means that the high frequency shear on the front half of the plate is out of phase with that of the

back half of the plate. This discussion relates only to the case of the high frequency shear commonly assumed to be a single shear along the length and thickness of the plate. Similar statements can be made if we consider the high frequency shear as being along the width and thickness.

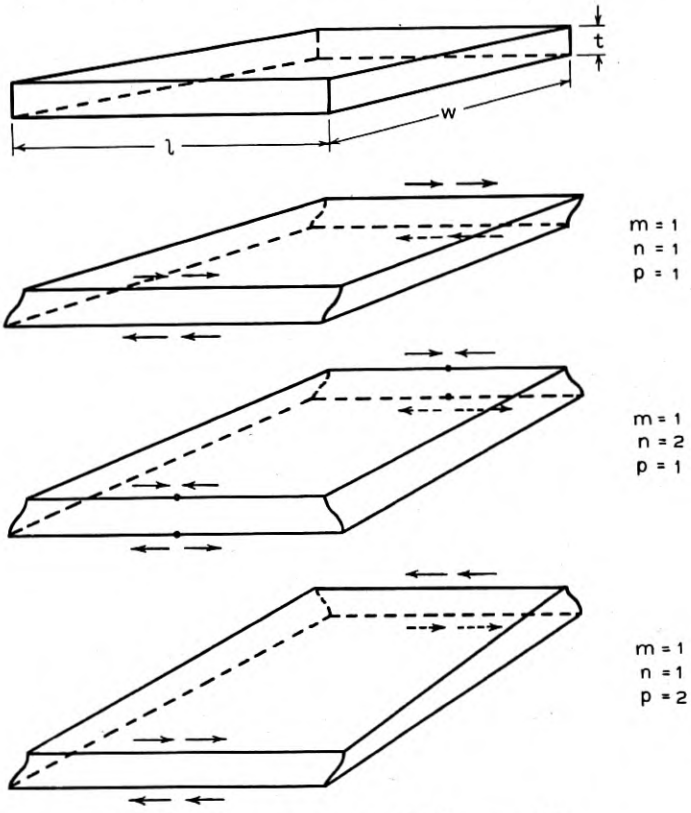

Fig. 6.5—Motion of a plate in high frequency shear.

6.24 *Type of Motion for Some Standard Filter and Oscillator Plates*

To get a more complete picture of the applications of the various types of motion, we will now take specific cases. The various crystals as commonly used for oscillators or filters are shown in Fig. 6.6. At the top of Fig. 6.6 are shown the various types of shear plates with their relative position with respect to the crystallographic axis.

The *AT* and *BT* plates are termed high frequency shear plates and the motion associated with them is that of a length-thickness shear as shown in Fig. 6.5. Their use is found for the control of radio frequency oscillators in

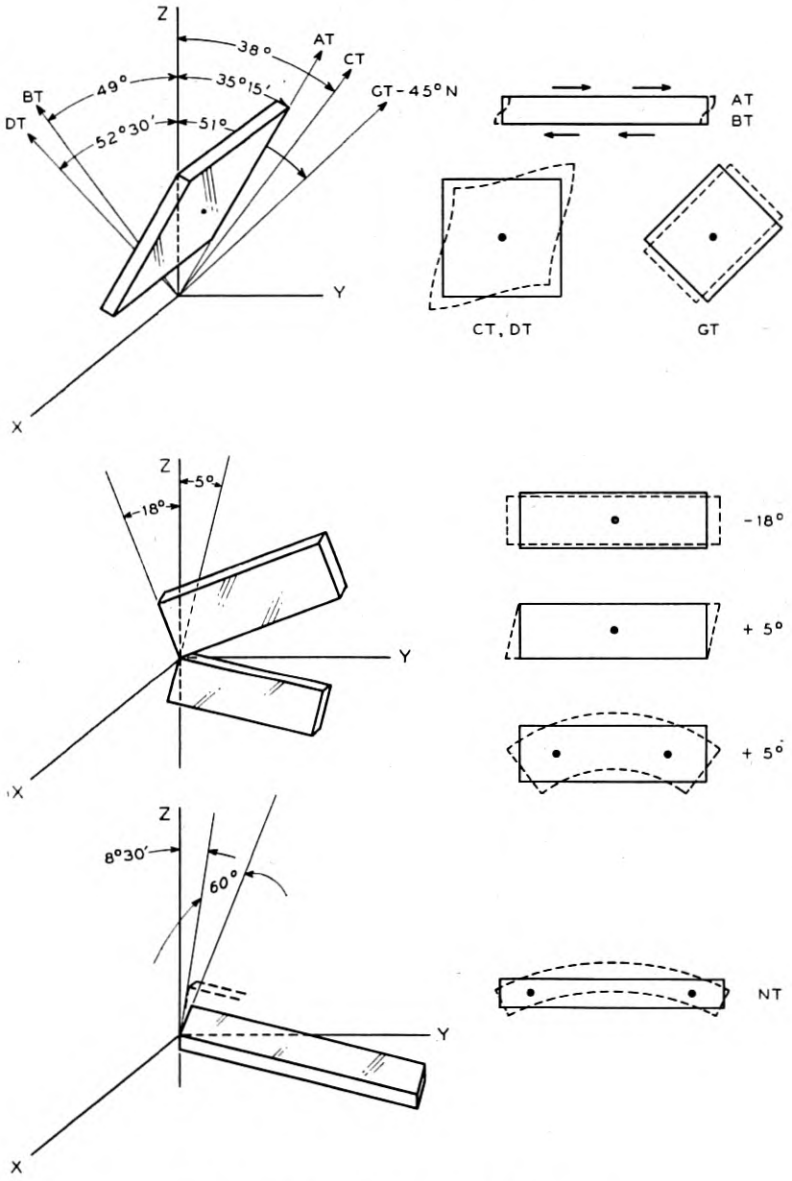

Fig. 6.6—Motions of typical cuts of quartz.

the range from 1 to 10 megacycles. The *AT* is most useful in the lower range and the *BT* in the upper range since it has a higher frequency constant.

Considerable use for the AT plate has been found for filters on pilot channels for the coaxial telephone system.

The CT and DT are analogous to the AT and BT but are termed low frequency shear plates. The motion associated with these cuts is that of a face shear as illustrated in Fig. 6.4. The CT and DT cuts are useful for both filter and oscillator applications in the frequency range from 60 kilocycles to 1000 kilocycles. Here again the DT would be most useful in the lower range and the CT the upper range due to the higher frequency constant for the CT cut.

The GT is similar to the CT except that it is rotated by 45° about the normal to the plate so that instead of a face shear type of motion there are two extensional modes similar to that shown in Fig. 6.3. These two modes are coupled to each other resulting in one of them having a zero temperature coefficient over a wide range of temperature. This crystal is most useful in the range from 100 kilocycles to 500 kilocycles for a primary standard of frequency and in filter networks having extreme phase requirements.

The filter plates commonly called the $-18°$ cut and 5° cut are shown with their relation to the crystallographic axes in the central part of Fig. 6.6. The $-18°$ cut commonly used in filters employs a simple extensional motion along its length with small coupling to an extensional motion along its width and practically zero coupling to a face shear type of motion. Since the width is usually the order of half the length these modes are not troublesome. The $+5°$ cut is useful in filter work because it has a low temperature coefficient and in spite of its strong coupling to the plate shear, it has been found quite useful in both its extensional mode and its flexure mode. The $-18°$ cut is used over the frequency range from 60 kilocycles to 300 kilocycles and forms the basic crystal used in the channel filters of the coaxial telephone system. When driven in flexure the 5° cut may be made to operate as low as 5 kilocycles and is used in oscillator and filter circuits.

The NT cut is shown at the bottom of Fig. 6.6 with its relation to the crystallographic axis. This is obtained by a rotation of $+8.5°$ about the X axis with a second rotation of $\pm$ 60° about the resulting Y' axis. The purpose of the second rotation is to give the shear modulus a positive coefficient. This modulus enters into the equation for the flexure frequency and therefore the effect of the second rotation is to change the temperature coefficient of the flexure mode from a negative value to zero. This crystal has been used to some extent as a low frequency oscillator. Its main purpose so far has been for the control of frequency modulation broadcast transmitters and for low frequency pilot channel filters.

Another crystal called the MT which is cut in a manner similar to the NT but with angles of 8.5° and 36° respectively has been used for filter work where an extensionally vibrating crystal of zero temperature coefficient is

required. The motion associated with this crystal is similar to that shown for the $+5°$ cut of Fig. 6.6. The low temperature coefficient is obtained through coupling to, and the effects of, a shear mode of positive temperature coefficient. Its use has been mainly for pilot channel filters of rather narrow frequency bands.

6.3 Frequency Equations for Flexurel, Extensional and Shear Motions

In determining the motion and resonant frequencies of a particular type of vibrating system it is customary to consider an isolated type of motion in order that the solution shall be in a simple enough form to be practical even though it may not be too accurate. The more accurate type of solution is often so complex that its use for practical solutions might be small. Since any solutions so far obtained are not complete in every detail, it is usually necessary to resort to experimentally determined frequencies in any case, and the solution can only be regarded as a guide to the complete result. In the following treatment it will be assumed that the frequency equations are given for isolated modes of motion and it will be later shown which of these forms are coupled and the effect of the coupling.

6.31 Flexural Resonant Frequencies

The simplest equation relating the resonant frequencies of a rod vibrating in flexure is given by[1]

$$f = \frac{m^2}{2\pi}\frac{k}{\ell^2}\nu \qquad 6.1$$

where ν = velocity of extensional propagation $= \sqrt{Y_0/\rho}$
k = radius of gyration of cross section
Y_0 = Youngs modulus
ℓ = length
$m \doteq (n + 1/2)\pi$ for free-free modes
$\doteq (n - 1/2)\pi$ for clamp-free modes ($n > 1$)
n = order of mode (1, 2, 3, etc.)

This equation holds only for the case of a long thin rod. Measurements of the resonant frequencies of a quartz crystal vibrating with both ends free has shown the above equation to be true where m is defined approximately as $(n + 1/2)\pi$ provided $\frac{nw}{\ell}$ is less than .1. For values greater than this the measured values are somewhat lower than that predicted. When the dimension in the direction of vibration is appreciable in comparison with the

[1] Rayleigh, Theory of Sound, Vol. 1, Chapter VIII.

length, Mason[2] has shown that it is necessary to consider the effects of rotary and lateral inertia. His solution leads to the same frequency equation as 6.1 but with a different evaluation of the factor m which is obtained from the transcendental equations

$$\left.\begin{aligned} \tan m X &= K \tanh mX' \text{ for even modes} \\ \tan m X &= -\frac{1}{K} \tanh mX' \text{ for odd modes} \end{aligned}\right\} \qquad 6.2$$

where

$$X = 1/2\left[\left(1 + \frac{m^4 k^4}{4\ell^4}\right)^{1/2} + \frac{m^2 k^2}{2\ell^2}\right]^{1/2}$$

$$X' = 1/2\left[\left(1 + \frac{m^4 k^4}{4\ell^4}\right)^{1/2} - \frac{m^2 k^2}{2\ell^2}\right]^{1/2}$$

$$K = \left[\left(1 + \frac{m^4 k^4}{4\ell^4}\right)^{1/2} + \frac{m^2 k^2}{2\ell^2}\right]^2 \frac{X'}{X}$$

Equation 6.2 holds only for the case of a rod free to vibrate on both ends. The case of a clamp-free rod is somewhat more complicated since it cannot be given by separate solutions for the even and odd modes. The interpretation of m given in equations 6.2 will result in the same value as before [$m = (n + \frac{1}{2})\pi$] for values of $\frac{nw}{\ell}$ less than .05 but decrease considerably for larger values and ultimately as the bar becomes wider the effects of rotary inertia result in the flexure frequency approaching the extensional mode as an asymptote. As stated before measurements on quartz bars vibrating in flexure departed from that predicted by the simple definition of m when the width of the bar was such that $\frac{nw}{\ell} > .1$. By using the value of m defined by equation 6.2 it is possible to predict the frequency for widths as great as $\frac{nw}{\ell} = .5$. For widths greater than this, experiment shows a frequency lower than that predicted by equation 6.2. This then leads one to believe that the effect of shear plays an important part in the flexure of bars with appreciable width. An investigation of the effect of shear on the flexure frequencies of beams has been made by Jacobsen[3] and his results lead to the same frequency equation as 6.1 and to the same transcendental equations derived by Mason (6.2) but with different values of X, X' and K to account for the shearing

[2] W. P. Mason, "Electromechanical Transducers and Wave Filters," Appendix A. D. Van Nostrand Company, Inc.

[3] *Jour. Applied Mechanics*, March 1938.

effect. These values are given by

$$X = \frac{1}{2}\left[\left(1 + \frac{m^4 k^4}{4\ell^4}\left(\frac{1}{c_{jj}s_{ii}} - 1\right)\right)^{\frac{1}{2}} + \frac{m^2 k^2}{2\ell^2}\left(\frac{1}{c_{jj}s_{ii}} + 1\right)\right]^{\frac{1}{2}}$$

$$X' = \frac{1}{2}\left[\left(1 + \frac{m^4 k^4}{4\ell^4}\left(\frac{1}{c_{jj}s_{ii}} - 1\right)\right)^{\frac{1}{2}} - \frac{m^2 k^2}{2\ell^2}\left(\frac{1}{c_{jj}s_{ij}} + 1\right)\right]^{\frac{1}{2}} \qquad 6.3$$

$$K = \left[\left(1 + \frac{m^4 k^4}{4\ell^4}\left(\frac{1}{c_{jj}s_{ii}} - 1\right)\right)^{\frac{1}{2}} + \frac{m^2 k^2}{2\ell^2}\left(\frac{1}{c_{jj}s_{ii}} - 1\right)\right]^2 \frac{X'}{X},$$

where c_{jj} is the shear constant in the plane of motion s_{ii} is the elastic constant in the direction of propagation. While it is true that these values will result in a lower value of m than those associated with equation 6.2 and hence fit the actual measured results more closely for bars wider than $\frac{nw}{\ell} = .5$, there is some doubt in the minds of various investigators as to the actual amount of correction necessary to apply to compensate for the shear. The solution of equation 6.2 using the constants of equation 6.3 is a lengthy process and could only be applied to a given orientation since the elastic constants vary with direction in quartz. While the results of Jacobsen's work are difficult to handle for intermediate values of $\frac{nw}{\ell}$ where the correction of rotary and lateral inertia do not fit the measured results it does imply that for large values of $\frac{nw}{\ell}$ that the flexure frequencies will be mainly a function of the length alone. Therefore when we are concerned with very high orders of flexure in plates such as the case of high frequency AT and BT shear crystals we may assume the interfering modes due to flexures will be essentially harmonic in nature. Restating the general problem of determining flexure frequencies in quartz rods or plates we may assume that the ratio of width to length is the controlling factor in deciding which method of attack is to be employed. For values of $\frac{nw}{\ell}$ less than .1 equation 6.1 will give quite accurate results. For values of $\frac{nw}{\ell}$ up to .5 equation 6.1, using the values of m determined by equation 6.2 will give satisfactory results. While the values of m determined by using equation 6.3 will give more accurate results for the range .4 to .6, it is not desirable to carry it further because, while 6.2 does take into consideration the effect of shear it does not account for coupling to the shear mode of motion. Hence for values of $\frac{nw}{\ell} > .6$

it is best to depend upon experimental measurements if accurate results are a factor.

6.32 *Extensional Frequencies*

The resonant frequencies of a bar vibrating along its length, commonly called an extensional mode of motion is derived quite easily from the wave equation in one dimension and is given by

$$f = \frac{n}{2\ell_i}\sqrt{\frac{1}{s_{ii}\rho}} \tag{6.4}$$

where ℓ = length
s_{ii} = elastic constant in the direction of propagation
ρ = density
n = 1, 2, 3, 4, etc.

This is the case when the length is the greatest dimension. When we consider extensional modes along the thickness of a plate, it can be shown that the c constants be employed to account for the lateral inertia in the two directions at right angles to the direction of propagation, (provided that the resulting motion is nearly along the thickness direction). Hence, for thin plates

$$f = \frac{n}{2\ell_i}\sqrt{\frac{c_{ii}}{\rho}} \tag{6.5}$$

As an example of the use of the above equation an X-cut bar vibrating along its length would result in a series of resonant frequencies defined by equation 6.4. An X-cut plate vibrating along its thickness would result in a series of frequencies defined by equation 6.5. Applying the appropriate constants

$$f_\ell = \frac{n}{2Y}\sqrt{\frac{10^{14}}{127.9 \times 2.65}}$$

$$= \frac{272}{Y(\text{cm})}\, n \quad \text{kilocycles} \tag{6.6}$$

and

$$f_t = \frac{n}{2X}\sqrt{\frac{86.05 \times 10^{10}}{2.65}}$$

$$= \frac{285}{X(\text{cm})}\, n \quad \text{kilocycles} \tag{6.7}$$

This shows that although Young's Modulus is the same in the two directions the resulting frequency constants are different because of the conditions at the boundaries.

6.33 *Shear Resonant Frequencies*

As shown in section 6.23 the low frequency face type shear mode results in a doubly infinite series of frequencies due to the manner in which the plate may break up into reversals of phase along its length and width. While a solution for the low frequency shear motion that satisfies the boundary condition of a free edge has not yet been accomplished, several approximate solutions for the frequencies are available. A modification of the equation developed by Mason[4] will give results which verify experimental data.

$$f = \frac{1}{2}\sqrt{\frac{1}{\rho s_{ij}}}\sqrt{\frac{m^2}{\ell^2} + k^2\frac{n^2}{w^2}} \qquad 6.8$$

where ρ = density
s_{ij} = shear modulus in ℓw plane
m, n = 1, 2, 3, etc.
ℓ = length of plate
w = width of plate

The value of k so far remains experimental and for low orders of m and n may be assumed unity. Its use is mainly for high orders of m and n where Young's modulus is different in the ℓ and w directions. Experimental data in the case of BT plates indicates that it should be 1.036 to account for the difference in velocity in the two directions. When m or n is large the velocity component, namely $\sqrt{\frac{1}{\rho s_{ij}}}$ should be replaced by $\sqrt{\frac{c_{ij}}{\rho}}$ for reasons explained for the extensional case. Equation 6.8 holds for the case of a plate vibrating in low frequency shear in regions where no highly coupled extensional or flexural resonant frequencies exist. As will be shown later, these regions are few. By assuming the frequencies are given by these equations and then applying the normal correction for coupled modes, a fairly accurate result will be obtained.

The high frequency case of a plate vibrating in shear is somewhat similar to the face shear or low frequency case with the exception that three dimensions must be considered since two are large compared to the third (the main frequency controlling dimension). An experimental formula for this case is given by

$$f = \frac{1}{2}\sqrt{\frac{c_{ij}}{\rho}}\sqrt{\frac{m^2}{t^2} + k\frac{n^2}{\ell^2} + k_1\frac{(p-1)^2}{w^2}} \qquad 6.9$$

where c_{ij} = shear modulus in plane of motion
ρ = density
ℓ, w, t = length, width and thickness

[4] "Electrical Wave Filters Employing Quartz Crystals as Elements," W. P. Mason, *B.S.T.J.* July, 1934.

m, n and p represent reversals of phase along the three directions and may be termed overtones. The values of k and k_1 are inserted to correct for the change in shear velocity resulting from a change in Young's modulus in the three directions. For most work with oscillator crystals where the length and width are large compared to the thickness, the following simplification of equation 6.9 is most useful.

$$f = \frac{m}{2t}\sqrt{\frac{c_{ij}}{\rho}} \qquad 6.10$$

When high frequency shear type crystals are used in connection with selective networks, it is necessary to make use of equation 6.9 to determine where the next possible pass regions will occur.

6.34 *Effects of Rotation About the Crystallographic Axes on the Resonant Frequencies and Coupling between Modes of Motion*

Several of the elastic constants have been used in equations expressing the resonant frequencies. Since most of the crystal cuts now in use are rotated at some particular angle about the X crystallographic axis, it is of interest to know the effect of this rotation upon the elastic constants since they determine the resonant frequencies and the coupling between certain of the modes of motion. The general stress-strain equations for an aeolotropic body are given in equation A.1 of Appendix A together with their definitions. In the case of quartz where the axes of the finished plate are aligned with the crystallographic axes the constants reduce to 7 and are shown in equation A.8. Examination of these equations shows that there are extensional and shearing strains resulting from dissimilar extensional and shearing stresses through the elastic constants s_{ij} and c_{ij}. This results in coupling between modes of motion where a so-called cross strain exists. These couplings may be made zero or small by proper orientation of the crystal plate about the X crystallographic axis. The mathematics of this operation is simplified by the use of matrix algebra[5]. Upon performing this operation a new set of elastic constants are obtained and are plotted graphically together with the piezoelectric constants on Fig. 6.7. From this figure we may see that the coupling resulting from the s'_{24} constant will be zero if the crystal plate is orientated by $-18.5°$ about X with respect to the crystallographic axis. This constant determines the coupling between the extensional mode along the length (Y' dimension) and the face shear mode ($Y'X'$ dimensions). This analysis resulted in the use of the $-18.5°$ cut in the channel filters of the coaxial system. Two other crystal cuts resulting in low coupling between different modes of motion are the AC and

[5] "The Mathematics of the Physical Properties of Crystals," W. L. Bond, *B.S.T.J.*, Jan. 1943.

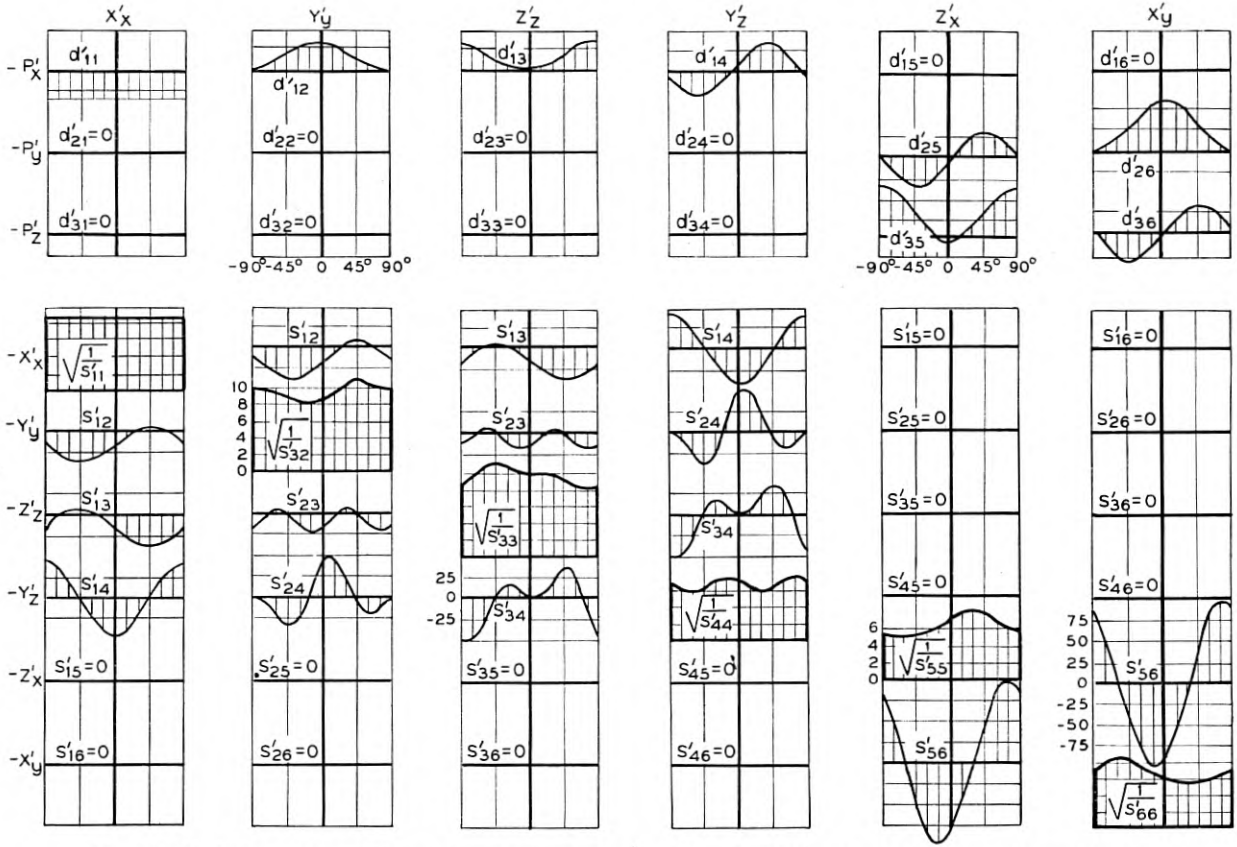

Fig. 6.7—Variation of piezo electric and elastic constants of quartz with rotation about the X crystallographic axis.

BC cuts. The s'_{56} constant determines the coupling between the face and thickness shear modes. As shown in Fig. 6.7 this constant passes through zero at two values, namely $+31°$ and $-59°$ and the resulting angles have been termed the *AC* and *BC* cuts. These angles are very close to the *AT* and *BT* cuts and hence they also possess the benefits of low coupling between modes. In addition to making the cross coupling constants zero, a rotation of the crystal plate with respect to the crystallographic axes also results in a change in the extensional and shear elastic constants. Notice that these pass through maxima and minima at the zero values for the cross coupling constants. This of course affects the resonant frequencies of isolated modes. Changes as great as 50% increase in frequency constants may be obtained by choosing the proper rotations. The equations relating the elastic constants as functions of orientation are given in appendix B for more complete use.

6.4 Coupling between Modes of Motion

As pointed out in the previous section, the frequency equation of a given mode of motion will give accurate results only in the case where the mode of motion is isolated. This is very rarely the case since most quartz crystals in common use are in the form of plates where the frequency determining dimension is not large in comparison with all other dimensions. Only in the case of a long thin rod vibrating in length-thickness flexure of the first order would this be true. It was also shown that the coupling between different modes of motion could be related to the mutual elastic constants (s_{ij} and c_{ij}) and that some of these could be made zero by the proper choice of orientation of the finished crystal plate. The elastic constants s_{ij} and c_{ij} only relate to the coupling between the extensionals, the shears and the extensional to the shear. For example s_{23} relates to the coupling between the extensional modes along the *Y* and *Z* axes, s_{56} relates to the coupling between the low and high frequency shear modes of a *Y* cut plate and s_{24} relates to the coupling between an extensional mode along the *Y* axis and a shear mode in the *YZ* plane. One other important coupling condition occurs and that is between the flexure and the shear modes. There is at present no mathematical theory relating this form of coupling except from simple assumptions that may be drawn from the fact that the shear modulus enters as a controlling factor in determining the frequency of a bar vibrating in flexure and from the similarity of the two types of motion near the boundaries. Since it is possible to have a definite coupling between extensional and shear modes there must be coupling between the extensional and flexure modes. It would be expected that it would be proportional to the coupling between the extensional and shear modes.

6.41 *Extensional to Shear and Extensional to Flexure Coupling*

The coupling between the extensional and shear motion can best be illustrated by taking the case of an X cut plate the length of which lies along the Y axis and the width along the Z axis. This is shown in Fig. 6.8 together with two other cases, one in which the plate is rotated about the X axis by $-18°$ and the other a similar rotation but $+18°$. Also in Fig. 6.8 is shown an enlarged view of the change in the elastic constants and frequency constants as a function of the rotation of the plate about the electric or X axis. For the case of an X cut plate the strains resulting from an applied exten-

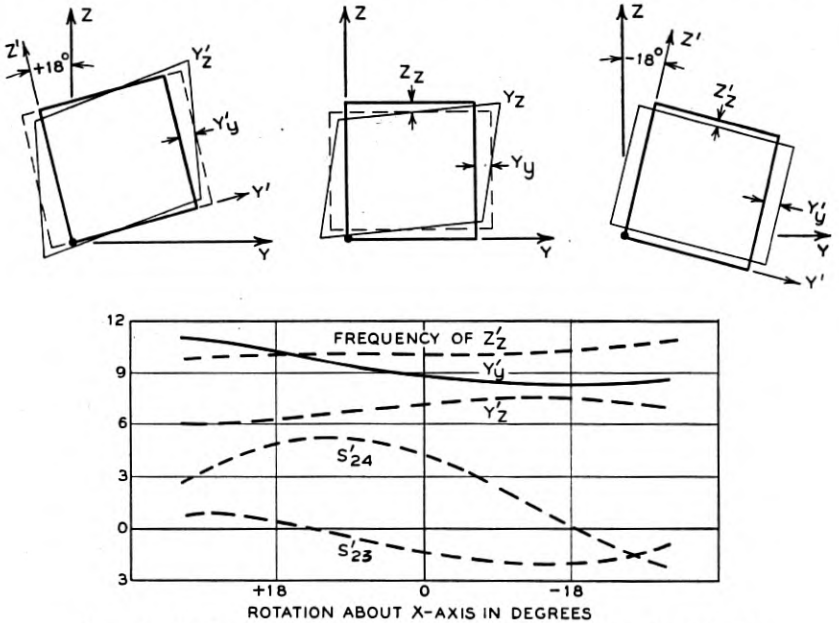

Fig. 6.8—Motion in an X cut plate for different orientation about the X crystallographic axis.

sional stress along the length according to equation A.8 would be

$$\begin{aligned} x_x &= s'_{12}\, Y_y \\ y_y &= s'_{22}\, Y_y \\ z_z &= s'_{23}\, Y_y \\ y_z &= s'_{24}\, Y_y \end{aligned} \qquad 6.11$$

where x_x is an extensional strain along the thickness
y_y " " " " " " length
z_z " " " " " " width
y_z " a shear strain in the length-width plane

If the plate is thin we may neglect the x_x strain as far as its effect on the resonant frequencies associated with the length and width are concerned. From the plot of the elastic constants on Fig. 6.8 we may determine the strains resulting from a stress along the length of an X cut plate for various orientations about the X axis. In addition to the expected extension along the length we have for a $+18°$ cut, a large amount of length-width or y_z' shear strain due to s_{24}' and very little width or z_z' strain. For the 0° cut there is also large length-width or y_z' shear strain and a width or z_z' strain. In the case of the $-18°$ cut the shear strain vanishes due to s_{24}' being zero, leaving in addition to the expected length or y_y' strain a width or z_z' strain. These relationships are more clearly shown if we plot the resonant frequencies resulting from the three modes of motion namely, the extensional modes along the length and width and the shear mode in the length-width plane

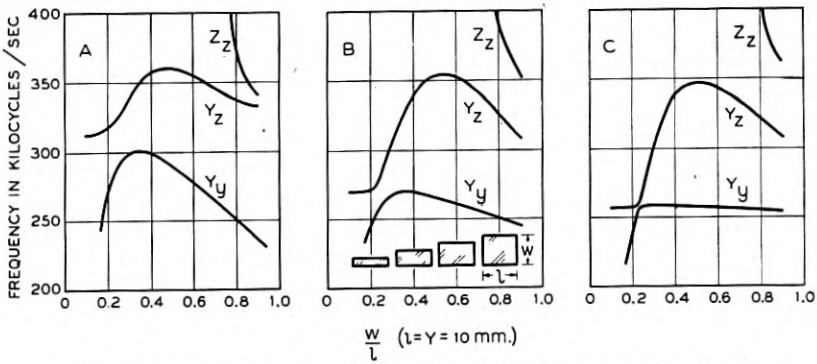

Fig. 6.9—Effect of rotation about the X axis on the resonant frequencies of an X cut plate.

A plot of measured resonances is shown in Fig. 6.9 for the above described three cases as a function of the change in width. The resonant frequencies for these three types of motion are given in section 6.3 as

$$f_{y_y'} = \frac{1}{2\ell}\sqrt{\frac{1}{\rho s_{22}'}}, \text{ extensional along } \ell \qquad 6.12$$

$$f_{z_z'} = \frac{1}{2w}\sqrt{\frac{1}{\rho s_{33}'}}, \text{ extensional along } w \qquad 6.13$$

$$f_{y_z'} = \frac{1}{2}\sqrt{\frac{1}{\rho s_{44}'}}\sqrt{\frac{1}{\ell^2}+\frac{1}{w^2}}, \text{ shear in } \ell w \text{ plane} \qquad 6.14$$

These equations specify only the uncoupled modes and do not take into consideration the effect of coupling to other modes of motion. In the case of Fig. 6.9 it is shown that when only the width is changed the extensional

mode along the length (the y'_y mode) is unaffected only in the case of the $-18°$ cut. The effect of coupling between the extensional and shear is clearly shown in the case of the 0° cut by the change in the length-extensional frequency. This is more pronounced in the $+18°$ case not because of more coupling but because the frequency constants of the two modes are more nearly alike as indicated in Fig. 6.8.

The mode of motion associated with the line intersecting the extensional y'_y mode is that due to the second length-width flexure mode. As mentioned before it is strongly coupled to the shear mode in the same plane. The coupling between this flexure and the extensional mode is directly related to the coupling between the shear and the extensional mode. This is borne out by Fig. 6.9, for in the case of the $-18°$ cut, s'_{24} is zero and as can be seen the change in frequency of the extensional mode is very slight even when the flexure mode is nearly identical in frequency.

We may state generally that the change in frequency of a particular mode of motion from that of its uncoupled state is dependant on two factors; the coupling to and the proximity to other forms of motion. This follows well established mathematical procedures but to solve the case just discussed would require the solution of a four mesh network with mutual impedances the values of some of which are at best only approximate. This will serve to illustrate that the use of formulae such as given in section 6.3 may be used more as a guide in establishing certain modes of motion rather than for accurate determinations of resonant frequencies.

6.42 *Flexure to Shear Coupling*

1. *Low Frequency Shear*

As previously indicated there is no simple means of mathematically determining the coupling between flexure and shear types of motion as there is between the extensional and extensional to shear modes. Here we must base our assumptions upon observed experimental evidence and simple reasoning. The relation between flexure motion and shear motion can be illustrated by the figures associated with Fig. 6.10. The forces that are necessary to produce flexure and shear motion are shown by arrows in Fig. 6.10. When the two arrows point toward each other, it indicates a compression and when the arrows point away from each other, it indicates tension. The diagrams on the left of Fig. 6.10 illustrate the conditions for flexure motion and the diagrams on the right indicate the conditions for shear motion. Notice that in the case of the first flexure and the second shear that the forces applied to the top and bottom of the plate are similar. Also in the case of the second flexure and third shear, they are similar. Here again we have certain similarities which in this case are important to remember.

The motion of the ends of the plate in the case of the first flexure are similar to those of the second shear. In the case of the second flexure the similarity is observed in the case of the third shear. The end motion in the case of the third shear is also the same in the case of the first or any odd shear. Likewise, the end motion of the first flexure is similar to the second shear or any even shear. We may then generalize and say that it is very likely that an odd order flexure would be coupled to an even shear; and also an even flexure would be coupled to an odd shear.

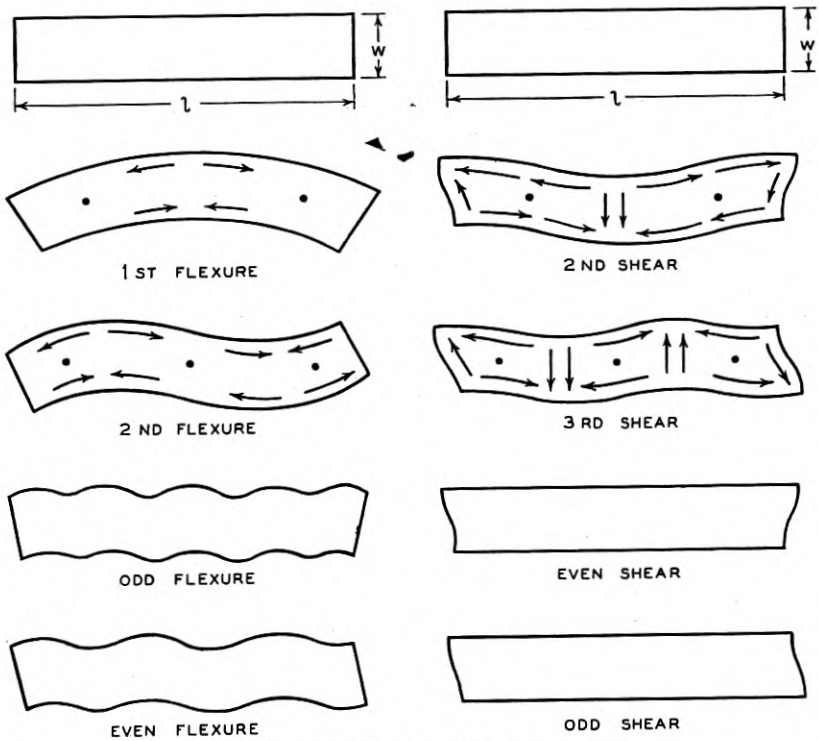

Fig. 6.10—Similarities in shear and flexure motions in a bar.

To illustrate the coupling between flexure and shear type motions, the frequencies of flexure and shear modes in a Z-cut quartz plate as shown in Fig. 6.11 have been measured. These measured frequencies are shown by the solid lines for various widths of the plate. It will be seen that there are no observed resonances following an unbroken continuous line to represent the shear frequency, but they are interrupted by several other frequencies which we must interpret as being various even modes of the flexure in the plane of the plate. It is clearly shown here that only even order flexures are

strongly coupled to the fundamental or odd shear. The strong coupling shown between the X_y shear and the second X_y flexure explains why the frequency equations given in section 6.3 for the frequency of flexure and

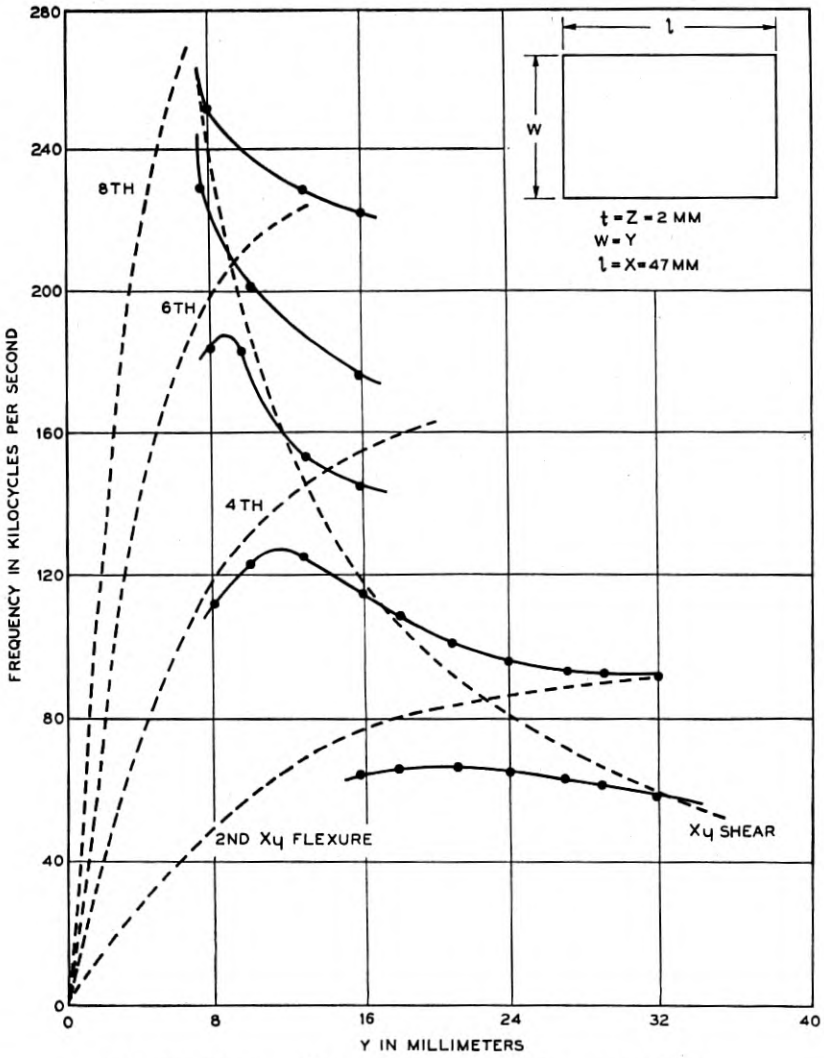

Fig. 6.11—Shear and flexure resonances in a Z-cut quartz plate.

shear modes will not give even approximate results if applied to this case for a square crystal. It will be shown later that if account is taken of coupling, the shear mode for a square crystal of this type may be more accu-

rately determined. Fig. 6.12 is a more detailed representation of the conditions shown broadly in Fig. 6.11 except in this case an AC-cut quartz plate was used and most of the observable resonant frequencies are shown

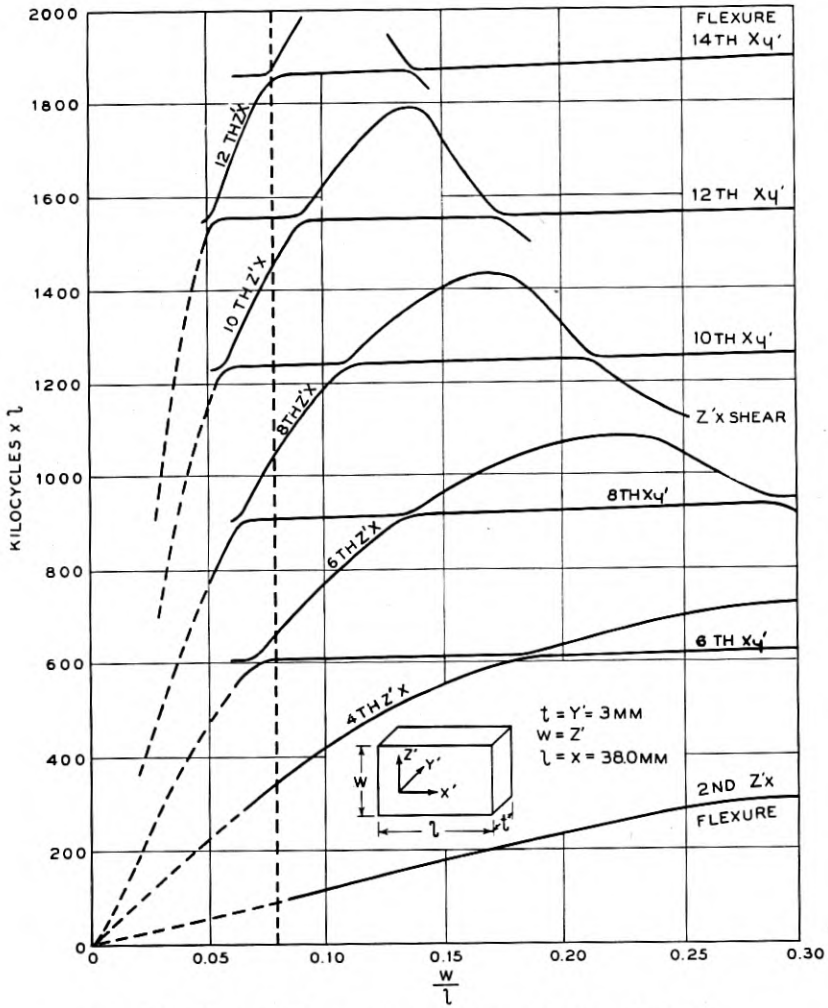

Fig. 6.12—Shear and flexure resonances in an AC-cut quartz plate.

for various values of $\frac{w}{\ell}$. The plate shear is labeled Z'_x shear and occurs at the frequencies predicted by equation 6.8 except in the regions where a flexure in the same plane exists. This is the type of motion shown in Fig. 6.4 for the case of $m = 1$, $n = 1$. It can be seen that as the difference in order of modes becomes greater the effect on the shear frequency is less

except where they are coexistant. We can then state generally that even though there is coupling between particular modes of motion, if the difference in order is great, the approximate frequencies may be computed as though they were isolated. This is more clearly shown in the case of thickness shear modes. The modes that are shown coupled to the face shear mode are Z'_x flexures propagated in the direction of the length or X axis. The lower orders can be shown to follow the general frequency equation discussed in section 6.3 but the higher orders for a given $\frac{w}{\ell}$, it will be noticed, are regularly spaced in frequency and show the effect of shear. The X'_y flexure modes determined by the length and thickness are shown as nearly horizontal lines since only the width was changed. Since these two groups of flexure modes are propagated in the same direction, it would be expected that the difference in frequency for the same ratio of dimension $\left(\text{i.e., } \frac{w}{\ell} = \frac{t}{\ell}\right)$ would be due to the differences of the shear coefficients in the two planes of motion. The vertical dotted line indicates the ratio of thickness to length. When the ratio of width to length is equal to this value it can be seen that the flexure modes in the width-length plane are in all cases higher than the same order flexures in the thickness-length plane. An examination of Fig. 6.7 shows that for an AC-cut crystal the shear modulus in the width-length plane $\left(\frac{1}{\sqrt{s'_{55}}}\right)$ is greater than that in the thickness-length plane $\left(\frac{1}{\sqrt{s'_{66}}}\right)$. This is in agreement with the observation made above. One other generality may be drawn from the experimental data shown in Fig. 6.12. The coupling between flexure modes and shear modes in planes at right angles to each other is very small in comparison with that between modes in the same plane.

As mentioned before the effect of coupling between modes of motion is greatest when the orders are more nearly similar. In this particular crystal this effect can be shown between the fundamental width-length Z'_x shear and the second order width-length Z'_x flexure. This is shown in Fig. 6.13 which is an extension of the data shown in Fig. 6.12 for a crystal nearly square and shows the frequency range covered only by the second flexure and the fundamental plate shear. A computation of the uncoupled second flexure mode propagated along the length and the first plate shear mode are shown by the solid lines f_f and f_s respectively. Inserting the appropriate constants the formulae of section 6.3 become

$$f_f = \frac{1}{2\pi}\sqrt{\frac{7.85 \times 10^{11}}{12 \times 2.65}}\, m^2 \frac{Z'}{X^2} \qquad 6.15$$

$$f_s = \frac{1}{2}\sqrt{\frac{71.8 \times 10^{10}}{2.65}}\sqrt{\frac{1}{X^2} + \frac{1}{Z'^2}} \qquad 6.16$$

In evaluating m, account was taken only of the rotary and lateral inertia so that some error is expected at the larger ratio of axes. The curve of flexure crosses the shear curve at $\frac{w}{\ell} = .76$, a condition which we know to be non-compatible since these two motions are coupled. From the theory of coupled circuits we can determine the displacement of two uncoupled frequencies as a result of the coupling, through the relation

$$f^2_{1,2} = \tfrac{1}{2}[f_s^2 + f_f^2 \pm \sqrt{(f_s^2 - f_f^2)^2 + 4k^2 f_s^2 f_f^2}] \qquad 6.17$$

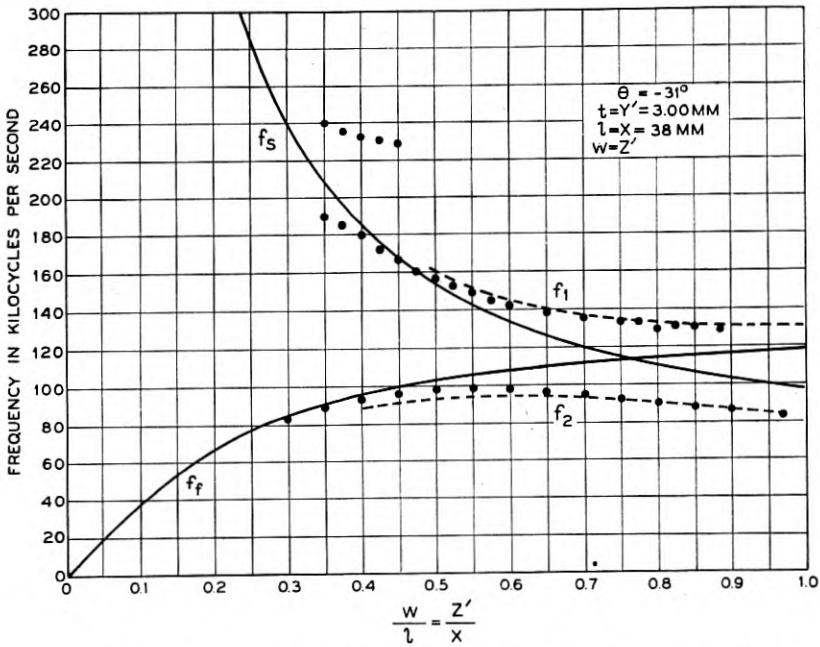

Fig. 6.13—Effect of coupling on the plate shear and the second flexure mode in an *AC*-cut quartz plate.

where f_s = uncoupled shear frequency,
f_f = " flexure "
k = coefficient of coupling.

The coefficient of coupling in this case may be defined as the ratio of the mutual to the square root of the self compliances of the two vibrating systems. As mentioned before no derivation has yet been made to indicate the relation between the coupling between these two forms of motion and the physical constants of the medium in which the vibration occurs. It is necessary to assume some coupling factor which will produce that observed

by experiment. Applying a coupling coefficient of 35% and computing the values of f_1 and f_2 from equation 6.17 the results are the dotted curves shown in Fig. 6.13. The observed points follow the computed values to a fair degree of accuracy for all frequencies below 180 kilocycles. Above this range there is a strong coupling to the fourth flexure and this would require separate consideration. Based upon these results the equation for the low frequency or face shear given in section 6.3 would not give the observed results for a nearly square plate because of the high coupling to the second flexure mode. For an approximately square plate, cut near the *AC*-cut the plate shear frequency including the effect of coupling would be given by

$$f = \frac{.849\sqrt{2}}{2d}\sqrt{\frac{1}{\rho s'_{55}}}, \qquad 6.18$$

where

$$d = \tfrac{1}{2}(X + Z')$$

and .849 is the factor resulting from the use of equation 6.17. For crystal cuts far different from the above it would be necessary to consider the flexure and shear as uncoupled and then apply equation 6.17 to determine the appropriate factor for square plates.

2. *High Frequency Shear*

The motion associated with flexure has been shown in Fig. 6.1 and in order to determine the frequency of higher order flexures, measurements were made on an *AC*-cut crystal. The results of these measurements are shown in Fig. 6.12. The first flexure motion to be expected with this crystal would be a flexure in the plane of the length and width. The various orders of these flexures are shown by the curved lines labeled second z'_x fourth, sixth, etc., all radiating from zero frequency (Primed values of z and y indicate that these are not crystallographic axes). The equation commonly determining the frequency of flexure states that the frequency should be proportional to the width and inversely proportional to the square of the length. If this were true, these curved lines representing the resonances of this type flexure shown on Fig. 6.12 would then be straight lines. Since the actual conditions show a wide departure from this, we must assume that this departure is due to rotary and lateral inertia and the effects of shear. It will be noticed that as we progressively increase the order of the harmonic, that the actual frequency spacing for a given value of $\frac{w}{\ell}$ is very nearly linear instead of a square law. This point is more clearly seen when we examine the frequency of higher orders of the flexures in the length thickness or xy' plane. As shown on Fig. 6.12 these frequencies

labeled 6th x_v, etc., change very little and are nearly horizontal straight lines. Here again they appear to be simple harmonics of some common low frequency. Also it will be noted that the coupling between the z_x' flexures and the z_x' shear is quite appreciable and in general decreases as the difference in order of the two modes becomes greater. This plot of the various flexure frequencies tells us a great deal about the behavior of progressively higher order of flexure type motion. The important effect to be noticed is that for high orders, and a fixed ratio of $\frac{w}{\ell}$, the flexure may be treated as though it were harmonic so far as frequency is concerned. Some variations to this rule will be observed and special cases will be discussed. So far we have discussed the case of flexure modes of relatively low order. In the case of high frequency shear modes of motion, we would expect that the order of flexure which would interfere with this type of motion would be rather high.

Figure 6.14 shows a plot of these flexure modes as observed in an AT-cut plate. These are shown by dashed lines. The dots indicate actual measured resonances. This figure also shows the various other resonant frequencies observed in this type of plate as discussed in section 6.2. The solid lines labeled mnp represent the type of shear motion shown in Fig. 6.5. Here again we may observe certain statements made before with respect to the coupling between shear and flexure type motions. Notice in this case that the coupling between an even order flexure and an odd order shear is high and increases as the orders more nearly approach each other. For example, the 38th flexure mode is coupled to the fundamental shear labeled $m_1n_1p_1$ has very little coupling to the second order shear $m_1n_2p_1$, and again is strongly coupled to the third shear $m_1n_3p_1$ and correspondingly higher coupling to the fifth shear. When we speak of higher order shears, such as $n_2n_3n_5$, they are not higher order in the sense of harmonics, but do differ by a small amount in frequency. In the case of a plate where ℓ is not great compared to t, these differences will be greater.

In actual practice in the case of AT plates, we are usually concerned mainly with the fundamental high frequency shear and high even order flexures along the length. This case is shown in Fig. 6.15 which gives experimental results of measurements on actual AT plates. It will be noticed that the flexure frequencies show a rather regular displacement as the ratio of the length of the plate to its thickness is changed. In this case only the odd order modes of shear and the even modes of flexure are shown. It will be observed that as the ratio of the length to thickness decreases, the coupling between these modes is quite high. This some state of affairs is illustrated again in the case of the third harmonic of high frequency shear and is shown in Fig. 6.16. The near vertical dashed lines represent even order

flexure frequencies and the curve labeled m_3n_1 and the curve labeled m_3n_3 correspond to two different values of the high frequency shear near its commonly called third harmonic.

An examination of Figs. 6.14 and 6.15 indicates that a regular pattern is formed of the ratios of axes at which the high frequency shear and succes-

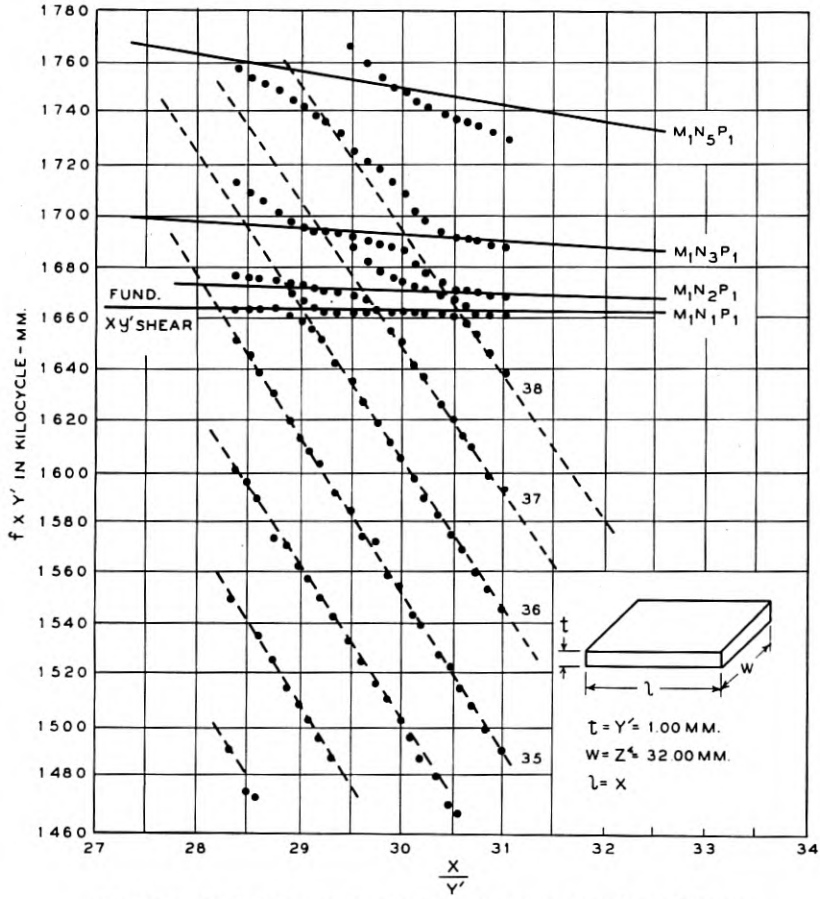

Fig. 6.14—High frequency flexure and shear resonances in an AT-cut quartz plate.

sive even orders of the length-thickness flexure coincide. Rather than define these points on the basis of specific ratios of axes it is more convenient to place them on a frequency basis. Therefore we may say that for a given size plate there will be specific frequencies at which some mode of the flexure motion along the length will be the same as the high frequency thickness

shear. For the case of AT plates experiment has shown these to be given by

$$f_{xf} = \frac{1338.4}{X} n_{xf}, \quad \text{kilocycles} \qquad 6.19$$

where X = length of X axis in millimeters,
n_{xf} = order of flexure along X axis
= 1, 2, 3, 4, etc.

In this equation as well as those of a similar nature to follow it is assumed that the thickness is such as to result in the same frequency for the high

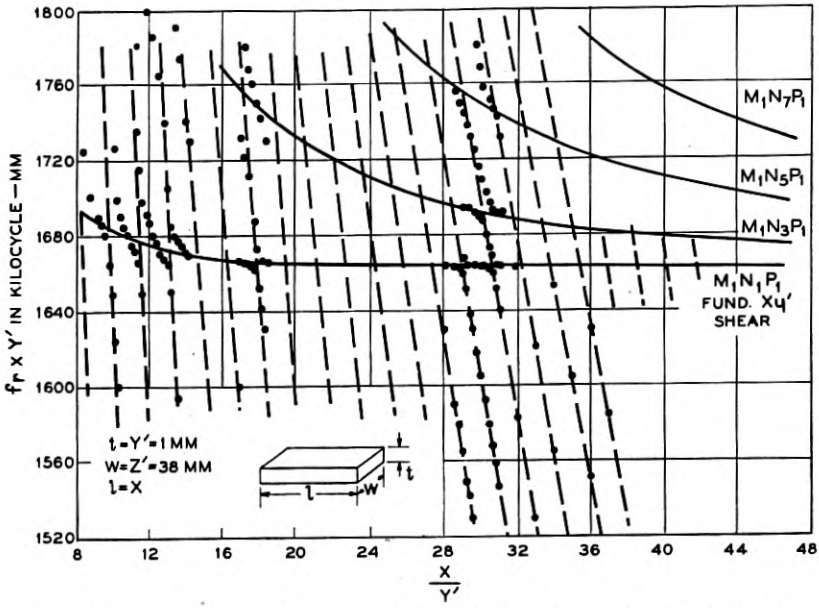

Fig. 6.15—High frequency flexure and shear resonances in an AT-cut quartz plate.

frequency X_y shear mode. As shown in Fig. 6.14 only the even orders are strongly coupled to the fundamental thickness shear.

The coupling between high even orders of the flexure along the X axis and the high frequency shear in the case of BT-cut plates is similar to that for AT-cut plates. Fig. 6.17 shows the various resonant frequencies observed in a BT-cut crystal as a result of changing the ratio of the length or X axis to the thickness or Y' axis. The curve m_1n_1 represents the high frequency X_y, shear. Curves m_1n_3, m_1n_5, m_1n_7 and m_1n_9 represent other X_y, shear modes as discussed in section 6.23 resulting from higher orders along the length or X axis. The dashed lines represent even order flexure modes along the X axis. The same regularity is observed here as in the case of the

AT-cut. When placed on a frequency rather than a ratio of axis basis the frequencies at which flexure modes along the X axis would coincide with the

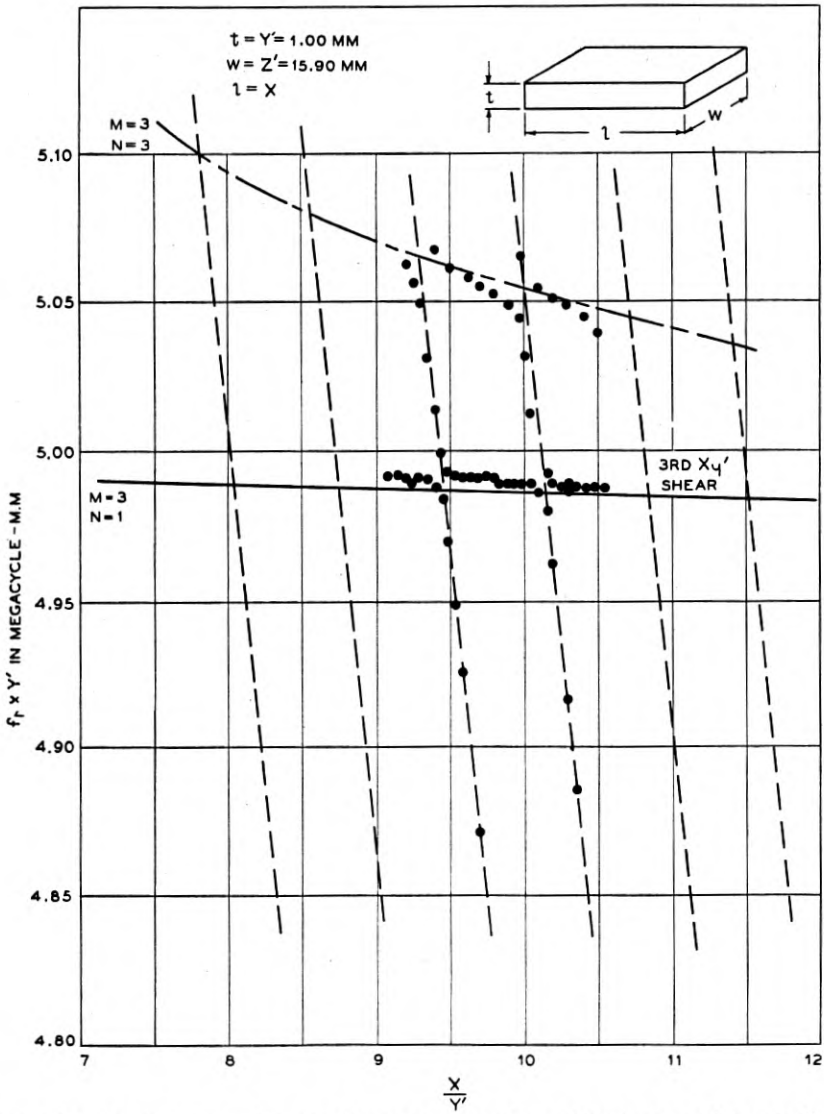

Fig. 6.16—High frequency flexure and shear resonances in an AT-cut quartz plate near the third harmonic shear mode.

fundamental $X_{y'}$ shear mode are experimentally given by

$$f_{xf} = \frac{1818}{X} n_{xf} \quad \text{kilocycles} \qquad 6.20$$

where X is given in millimeters. In this case it will be noticed also that only even order flexures are strongly coupled to the fundamental $X_{y'}$ shear.

The dependence of the flexure frequency on the shear coefficient can be seen from these two cases. The direction of propagation is the same in both cases (along the X axis) but the direction of particle motion is nearly at right angles. It would be expected then that the frequency constant would be highest for the case of the highest shear coefficient. Examination of equa-

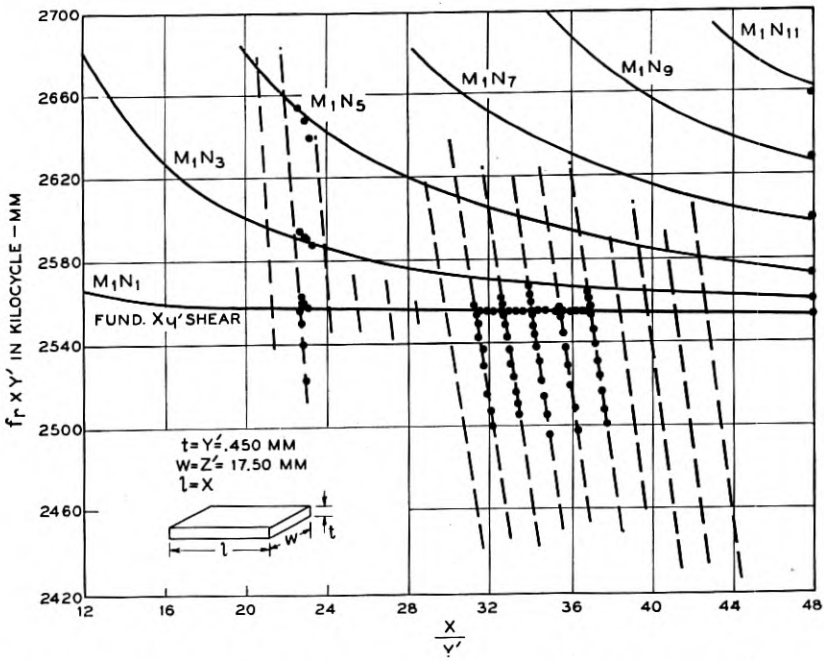

Fig. 6.17—High frequency flexure and shear resonances in a BT-cut quartz plate.

tions 6.19 and 6.20 shows this to be true. In addition, the change in the frequency constant is about the order of magnitude of the change in the shear modulus in the respective planes of motion.

6.43 *Coupling between Low Frequency Shear and High Frequency Shear*

From an examination of Fig. 6.7 it can be seen that the coupling between the low frequency shear (Z'_x) and the high frequency shear Xy' is related by the s'_{56} constant. In the AC and BC-cuts this reduces to zero but for the AT and BT-cuts it has a finite small value. According to section 6.3 the frequencies of the plate shear modes are given by equation 6.8 but this holds only for the case where m and n are small. When the third dimension

becomes appreciable in comparison with a half wave length along w or l it becomes necessary to use the c constants. When considering high orders of the low frequency shear equation 6.8 is modified to

$$f = \frac{1}{2}\sqrt{\frac{c_{ij}}{\rho}}\sqrt{\frac{m^2}{l^2} + k^2\frac{n^2}{w^2}} \qquad 6.21$$

Equation 6.21 shows that high orders of the low frequency or plate shear are dependent upon both the length and width dimensions and it might be assumed that this would lead to very complicated results in so far as analysis of experimental data is concerned. The coupling between these modes and the high frequency shear is a result of coupling in the mechanical as well as

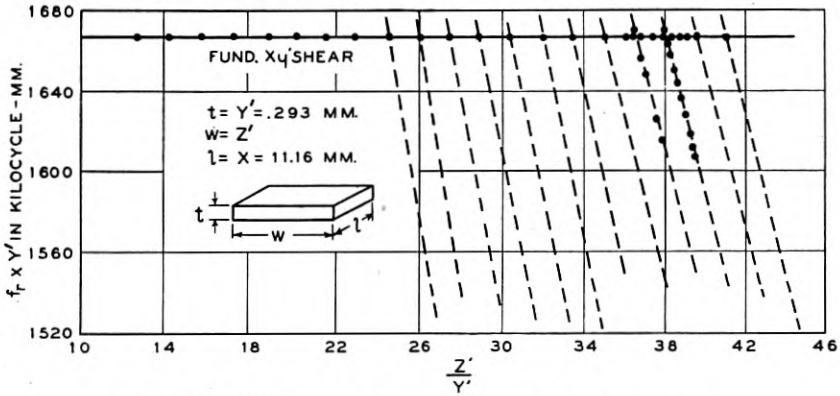

Fig. 6.18—High frequency shear resonances in an AT-cut plate.

the electrical systems. The strongest coupling with reference to the length axis would then be for high odd orders of m and unity for n with successively smaller coupling for higher orders for n if the driving potential extends over the complete surface of the crystal. In a similar manner when considering high orders of plate shear along the width axis the highest coupling will result from unit order for m. Based on these assumptions then to a first approximation we can assume these modes to be functions of length and width alone. Equation 6.21 then reduces to

$$f_{l_s} = \frac{1}{2}\sqrt{\frac{c_{ij}}{\rho}}\frac{n_{l_s}}{1} \qquad 6.22$$

$$f_{w_s} = \frac{k}{2}\sqrt{\frac{c_{ij}}{\rho}}\frac{n_{w_s}}{w} \qquad 6.23$$

where n_{sl} = order of shear mode along l axis,
n_{sw} = order of shear mode along w axis.

These modes have been measured in AT and BT-cut crystals. Fig. 6.18 shows the points at which these modes intersect the fundamental high frequency shear mode in AT-cut plates. This is the case for high orders along the Z' or width axis. A similar set of resonances can be shown to exist when the X or length axis is varied. Experiment has shown that these frequencies of coincidence between high order plate shear modes and the fundamental high frequency $X_{y'}$ shear mode for the case of AT-cut plates is given by

$$f_{xs} = \frac{254.2}{X} n_{xs} \text{ kilocycles} \qquad 6.24$$

$$f_{z's} = \frac{254.0}{Z'} n_{z's} \text{ kilocycles} \qquad 6.25$$

where X and Z' are given in centimeters. Only odd orders are strongly coupled if the crystal plate has a symmetrical contour with respect to an applied equipotential electrode. Upon substitution of the value of c'_{55} for an AT-cut crystal in equation 6.22 there results

$$f_s \times \ell = \frac{1}{2}\sqrt{\frac{c'_{55}}{\rho}} = \frac{1}{2}\sqrt{\frac{67 \times 10^{10}}{2.65}} = 251.0 \text{ kilocycle} - \text{cm}. \qquad 6.26$$

which is within 1 per cent of that found experimentally. Since Young's modulus is nearly the same along the X and Z' axis the value of k in equation 6.23 is essentially unity. Fig. 6.19 shows measured values of high order Z'_x shear modes near the high frequency $X_{y'}$ shear mode in a BT-cut crystal for various values of the width or Z' axis. More detailed measurements have been made of the high order Z'_x plate shear modes in BT-cut plates along the X axis. Fig. 6.20 shows both the shear and flexure modes along the X axis near the vicinity of the high frequency $X_{y'}$ shear mode. Since the frequency constant for the Z'_x shear modes is different from that for the $X_{y'}$ flexures there are regions where, if no coupling existed, all three modes would be at the same frequency. It is obvious from Fig. 6.20 that this is not the case. Therefore, we must assume that not only are the high order Z'_x shears and $X_{y'}$ flexures coupled to the high frequency $X_{y'}$ shear but that they are coupled to each other.

While it is difficult to see from Fig. 6.20 the relative coupling of flexures to the $X_{y'}$ shear, experiment has shown the flexure modes along X to have the greater coupling to the $X_{y'}$ shear. This is true when the ratio $\frac{X}{Y'}$ is such that the flexure modes along X and high order Z'_x shear modes along X have their maximum separation. When these modes approach each other

and the $X_{y'}$ shear such as is shown in Fig. 6.21 at $\frac{X}{Y'} = 31.35$ the relative coupling of each to the $X_{y'}$ shear is about equal. This arises from the fact that the mutual coupling between them increases the apparent coupling

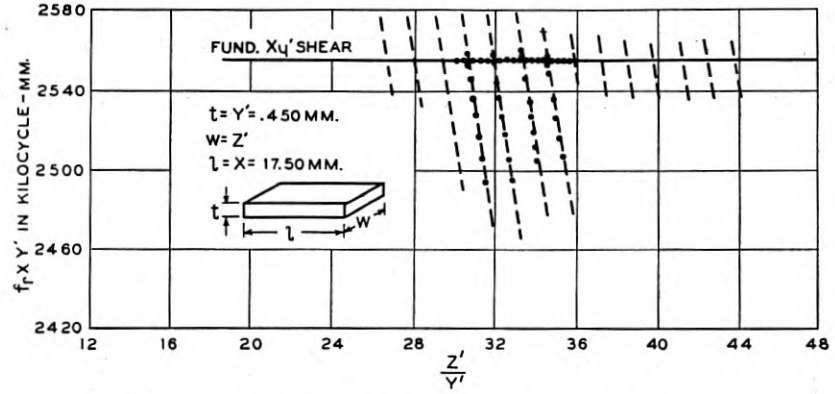

Fig. 6.19—High frequency shear resonances in a *BT*-cut plate.

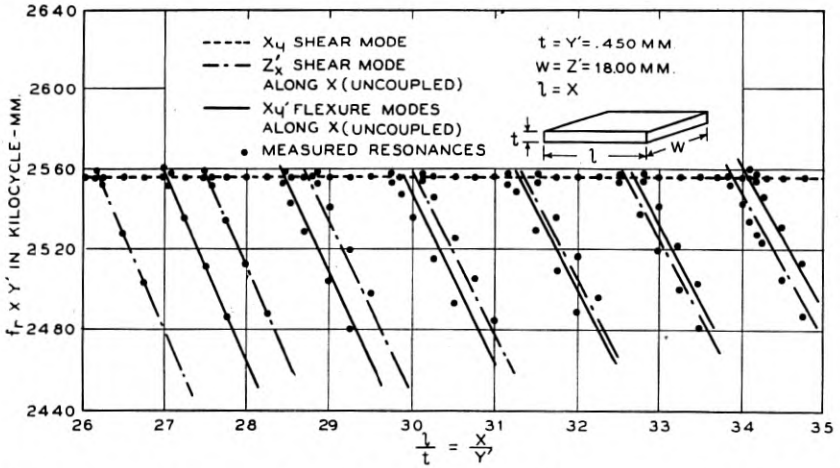

Fig. 6.20—High frequency thickness shear and flexure and shear resonances along the X axis in a *BT*-cut quartz plate.

between the $X_{y'}$ shear and high orders of Z'_x shear along X. From this it would appear advisable to avoid such regions in the dimensioning of crystals for oscillator use over wide temperature ranges. Determination of the flexure as well as high order Z'_x shears then must be made in regions where

they are spaced so that the effect of coupling between them will not influence the frequency constant that is determined experimentally. These regions have been investigated and the result for the flexure modes is that shown

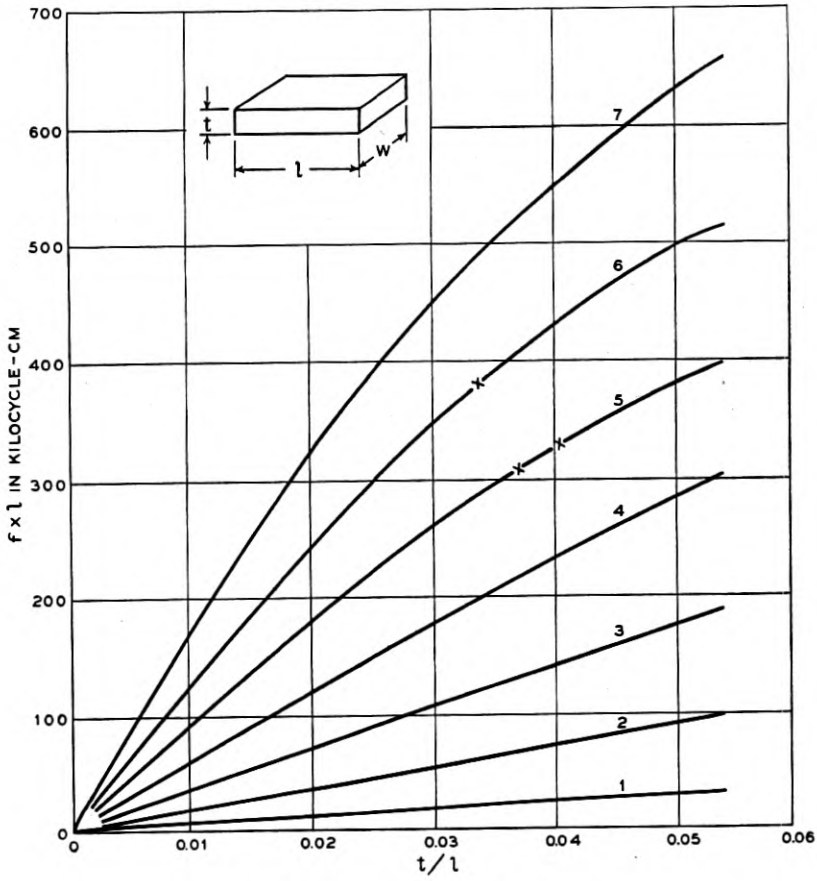

Fig. 6.21—Flexure resonances in a *GT*-cut quartz plate.

in equation 6.20. From Fig. 6.19 the high order Z'_x shears along Z' will be coincident with the high frequency $X_{y'}$ shear at frequencies given by

$$f_{z's} = \frac{166.45}{Z'} n_{z's} \quad \text{kilocycles} \qquad 6.27$$

From Fig. 6.20 high orders of the same Z'_x shear along X will be coincident with the high frequency X'_y shear at frequencies given by

$$f_{xs} = \frac{163.514}{X} n_{xs} \text{ kilocycles} \qquad 6.28$$

Upon substitution of the value of c'_{55} for a BT-cut in equation 6.22 there results

$$f_{t_s} \times \ell = \frac{1}{2}\sqrt{\frac{c'_{55}}{\rho}} = \frac{1}{2}\sqrt{\frac{30.3 \times 10^{10}}{2.65}} = 169.0 \text{ kilocycles} - \text{cm.}$$

which is 3.3% greater than that observed in equation 6.28 and 1.6% greater than that shown in equation 6.27. The apparent difference in the observed shear modulus in the X and Z' directions for the BT-cut can be explained from the fact that Young's modulus is quite different in the two directions for the BT-cut while it is nearly the same for the AT-cut as verified by equation 6.24 and 6.25.

From the discussion in this section it can be seen that a single theory that would relate all the now known resonances in quartz plates together with the effects of coupling would be prodigous indeed. In order to reduce the design of quartz plates to a simple engineering basis it is necessary to take specific examples and investigate the region in the vicinity of the frequency to be used based on general theory and then apply approximations that fit the specific cases.

6.5 Methods for Obtaining Isolated Modes of Motion

6.51 *GT Type Crystals*

In the case of GT type crystals the modes that cause the greatest concern are flexure modes in the two planes of the length and thickness and the width and thickness. The desired mode is that of an extensional mode along the width. To produce a low temperature coefficient it is also necessary that this mode be coupled to an extensional mode along the length, a fixed frequency difference from it. Therefore it will be necessary to prevent flexure modes from occurring at either of these two frequencies. Fig. 6.21 shows the frequency of various flexure modes that would be observed in GT-cut plates for different ratios of thickness to length. In the case of the GT-cut the elastic constants in the length and width directions are the same and therefore it is only necessary to determine the flexures in one plane to get a determination in both. From the plot of frequencies shown in Fig. 6.21, it would be very easy to determine the proper thickness for any given GT plate. Since in all practical cases there is a definite relation between the length and width of this type of plate, it would be necessary to examine the flexures in these two directions as a function of the change in thickness.

Fig. 6.22 shows a plot of this for the case of a *GT* crystal designed to operate at 164 kilocycles. All the information shown in this figure is obtained directly from Fig. 6.21. Since a change in thickness will not have any effect upon the length and width extensional modes of vibration and only

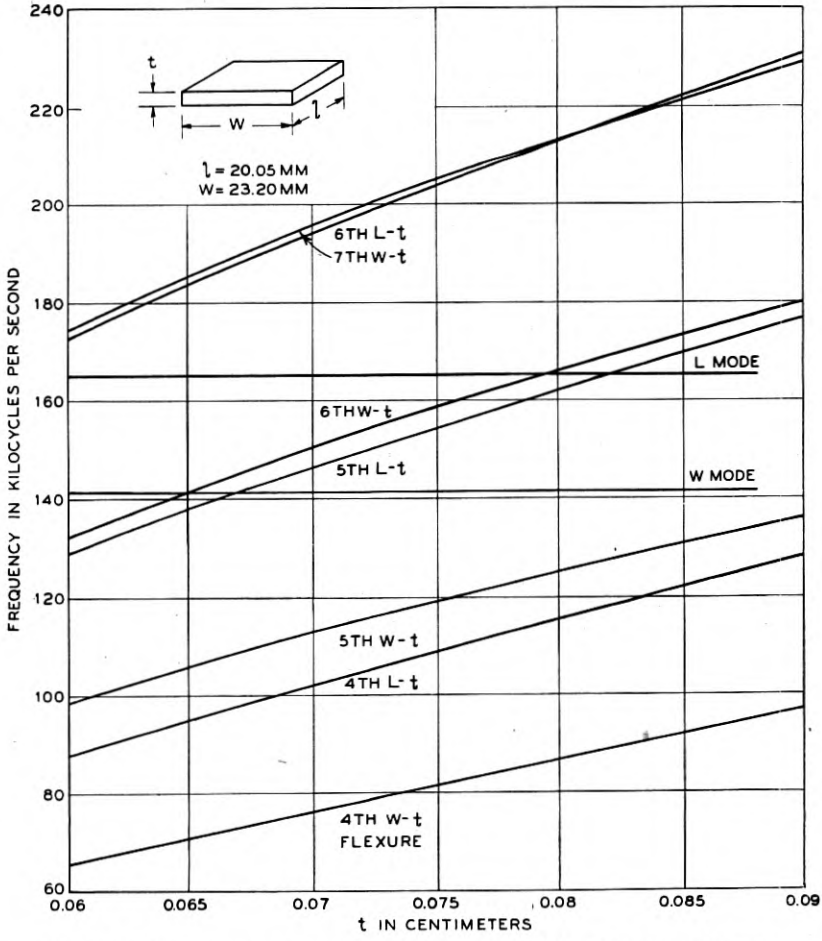

Fig. 6.22—Flexure and extensional resonances in a 164 kc *GT*-cut quartz plate.

changes the flexure frequencies, it would be reasonable to suppose that some thickness could be obtained where no flexure along the length or width would be of the same frequency as the length or width extensional mode. Examining the curves of Fig. 6.22, we find that a thickness of .06 cm., .075 cm. or .085 cm. would meet these conditions.

6.52 *BT Type Crystals*

As discussed in Section 6.4 the modes showing the greatest coupling to the high frequency thickness shear are of two types: high orders of $X_{y'}$ flexure propagated along the X axis and high order Z'_x shears along the X and Z' axes independently. Complex orders of the flexure and plate shear as illustrated in Fig. 6.2 and Fig. 6.4 do cause considerable difficulty and their analysis calls for special treatment and is not within the scope of this text. For the case of the BT-cut the three primary interfering series of modes are given by

$$f_{xf} = \frac{181.8}{X} n_{xf} \text{ kilocycles}$$

$$f_{xs} = \frac{163.514}{X} n_{xs} \text{ kilocycles} \qquad 6.30$$

$$f_{z's} = \frac{166.45}{Z'} n_{z's} \text{ kilocycles}$$

where X and Z' are given in centimeters and f_{xf} is the frequency at which integral orders of flexure modes along the X axis would coincide with the high frequency thickness shear mode. In a similar manner f_{xs} and $f_{z's}$ relate the same conditions for integral orders of the plate shear modes. These equations are true only in the case where the thickness is of such a value as to place the high frequency thickness shear mode at the same frequency as the computed interfering mode. In most practical cases for oscillator use the electric field is applied to the crystal by means of a flat electrode on each side of the crystal plate. Under this condition only odd order $X_{y'}$ shear modes along the X axis are excited and hence the strongest couplings to the $X_{y'}$ flexure modes will be only for even order values of n_{xf} in equation 6.30. In a similar manner the greatest interference between the $X_{y'}$ shear mode and high orders of the Z'_x shear modes along both X and Z' will occur for odd orders. Therefore the strongest interference from these modes will occur only for odd integers of n_{xs} and $n_{z's}$ in equation 6.30. These assumptions of only even flexures and odd shears showing appreciable coupling are based upon a crystal plate cut precisely along its proper axis and of uniform contour assembled in a holder using electrodes of uniform air gap. Deviations from these conditions will of course alter the ideal results dependent upon the amount and type of deviation.

The relationships shown in equation 6.30 may be more clearly seen when plotted graphically. Assuming a BT-cut crystal plate 1 centimeter square we may determine the frequencies at which an interfering mode will coincide with the high frequency $X_{y'}$ shear by assigning even integers to n_{xf} and odd

integers to n_{xs} and $n_{z's}$. Fig. 6.23 shows a plot of these three types of interfering modes on a folded frequency scale covering the range from 5 to 15 megacycles for a plate 1 centimeter square. Each abscissae covers a range of one megacycle with dots at three levels. The first level shows the frequencies at which successive even orders of flexure along the X axis occurs. The second level shows successive odd Z'_x shear modes along X and the third level successive odd Z'_x shear modes along Z'. The circles shown on the three levels indicate the results of actual measurements on BT-cut crystals as resonating elements. It will be noticed that the circles and dots coincide for most frequencies, the regions of departure occur only when a high order shear mode and a high order flexure mode along the X axis approach each

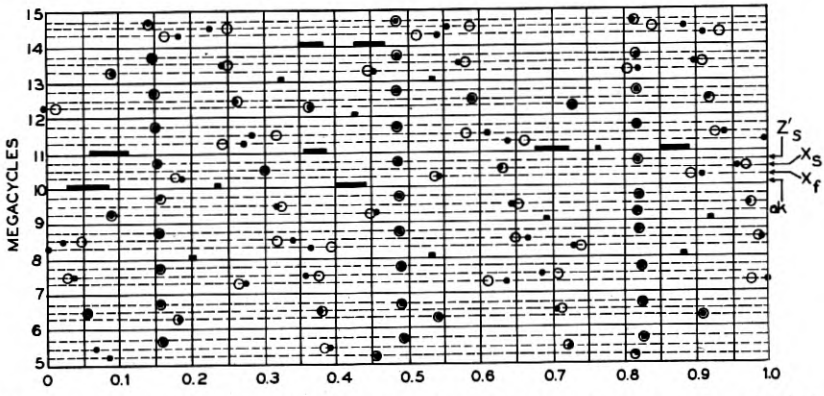

Fig. 6.23—Frequencies at which the Z'_x shear along X, the Z'_x shear along Z' and the X_y flexure along X coincide with the high frequency X_y shear in BT-cut crystals.

other in frequency. The reason for this is obvious from the previous discussion on the coupling between flexure and shear modes of motion.

The chart of Fig. 6.23 is of course not limited to a crystal 1 centimeter square or for that matter even a square crystal. In reality it relates the product of the frequency and X and Z' dimensions. For example a flexure mode interferes with the high frequency shear mode at a frequency of 9.45 megacycles for a plate with X dimension equal to 1 centimeter. If the X and Y' dimensions were doubled the same situation would exist at one half the frequency. In determining the dimensions for a crystal at a given frequency we know that the product of the frequency and X dimensions as well as Z' dimension must not result in a frequency close to those given by the circles of Fig. 6.23. In addition other interfering modes as previously mentioned must be avoided. These at present may be determined experimentally by choosing regions on the chart clear of the known flexure and shear modes.

On the abscissae are shown certain discreet frequencies as well as frequency ranges which have been found to result in crystal units having no serious dips in activity over a wide range in temperature. These are for square crystals in the 18 millimeter size range and have been obtained by Mr. G. M. Thurston of the Bell Laboratories and Mr. F. W. Schramm of the Western Electric Company. It will be noted that no so-called ok regions have been found at the frequencies of the three principal coupled modes.

While the use of the chart shown in Fig. 6.23 will often lead directly to the proper X and Z' dimensions for a given oscillator it cannot be overemphasized that only the three principal interfering modes are shown and only the odd orders for the shears and only the even orders for the flexure modes. Since the even order shear modes are excited due to slight variations which would produce wedge shaped air gaps or quartz blanks, it is advisable to avoid these regions also. Complex combinations of the three principal modes as shown in Figs. 6.2 and 6.4 are also driven. Therefore when it is necessary to produce a crystal unit possessing the highest activity for a given area of quartz plate over an extended temperature range it is necessary to scan the supposed desirable regions shown in Fig. 6.23 by complete measurements on finished units of a given size and varying frequency or of constant frequency and varying size. As an illustration the region shown in Fig. 6.23 between 10.025 and 10.080 megacycles was determined in this manner with the use of crystal plates approximately 18 millimeters square. The use of crystals with other than square dimensions could undoubtedly have increased the range of this region but their use is undesirable from a manufacturing standpoint. Assuming that the electrodes and crystal holder permit a variation in size of the quartz plate from 17.20 millimeters to 18.20 millimeters this approved region will immediately specify the dimensions of crystals to cover the frequency range from 5508 to 5727 kilocycles. This also assumes crystal blanks cut to precise orientations with controlled contours and electrodes of uniform flatness and constant airgap. While the theory would indicate that the frequency range given above could be expanded to considerably higher values by utilizing a smaller crystal blank this has not been proven so far since most crystals produced by the Western Electric Company require large area plates to meet high activity requirements.

As an illustration of the effect on the behavior of oscillators of changing the X and Z' dimensions of BT-cut quartz plates measurements have been made of the activity, in a conventional tuned plate circuit with the crystal connected between grid and cathode of quartz plates of constant thickness and varying X and Z' dimensions. Fig. 6.24 shows the effect of changing the X dimension of a quartz plate on its activity as an oscillator. By taking the product of the frequency and dimension we can determine the dimen-

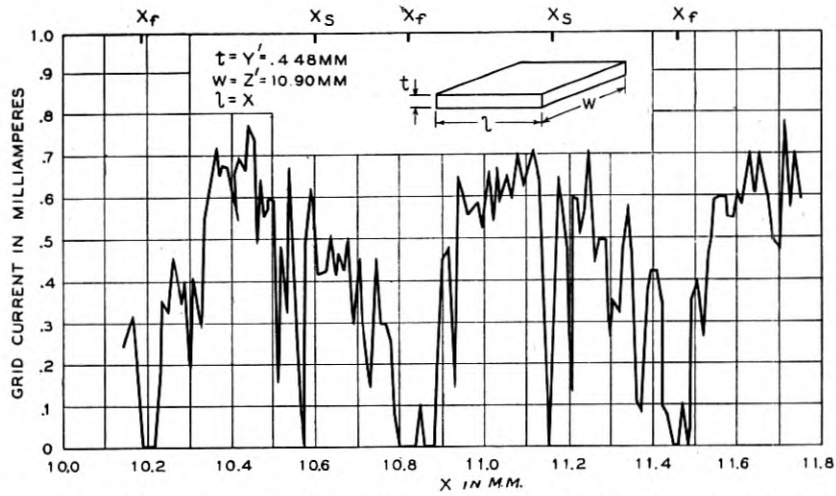

Fig. 6.24—Effect of change in X dimension on the activity of a BT-cut quartz plate in an oscillating circuit.

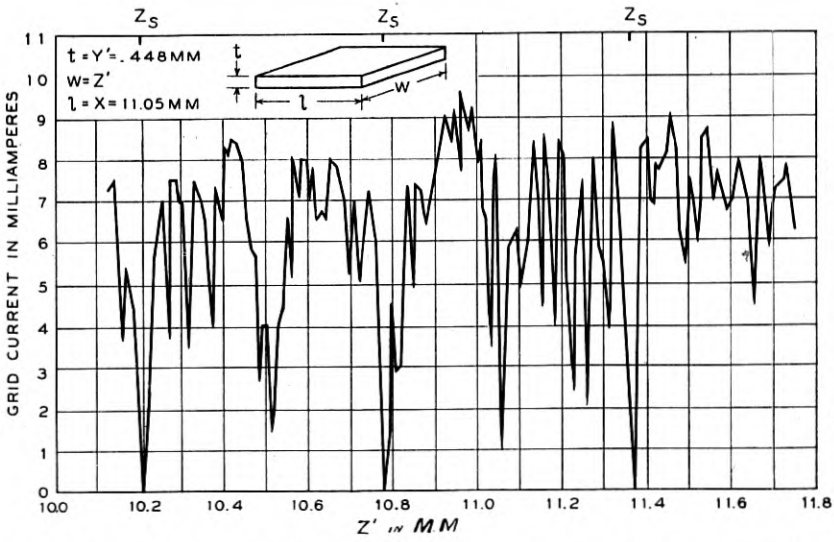

Fig. 6.25—Effect of change in Z' dimension on the activity of a BT-cut quartz plate in an oscillating circuit.

sions from Fig. 6.27 for this case where the $X_{y'}$ flexures and Z'_x shears will interfere to produce poor characteristics. These are shown in Fig. 6.24 for flexure modes as X_f and for the shear modes as X_s and do in general cor-

respond to the dimensions resulting in low or no activity. This illustrates quite clearly the necessity for grinding the edges of plates not dimensioned for a specific frequency. Fig. 6.25 shows the same conditions when only the Z' dimension is changed. In this case the dimensions shown at regular intervals as Z_s were derived from Fig. 6.25 as before and correspond to the zero activity dimensions found experimentally. It will be noticed that low activity regions are found halfway between the dimensions designated as Z_s. These correspond to even orders of the Z'_x shear and are the result of a slight wedge in the airgap. This was intentional to show the existence of this condition.

Figures 6.24 and 6.25 show the necessity for avoiding certain dimensions for oscillator plates at specific frequencies. This can be accomplished by

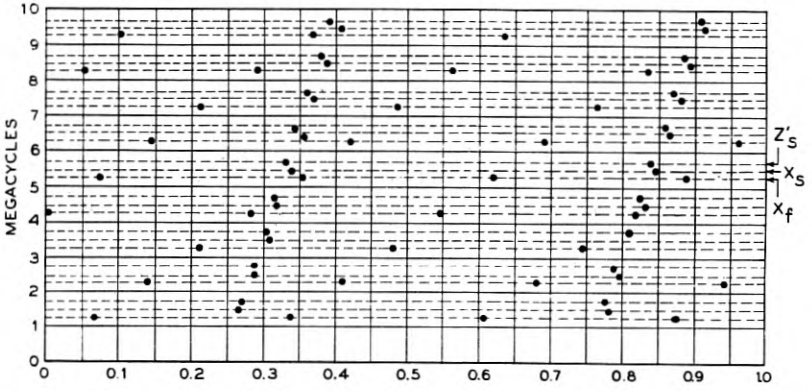

Fig. 6.26—Frequencies at which the Z'_x shear along X, the Z'_x shear along Z' and the X_y flexure along X coincide with the high frequency X_y shear in AT-cut plates.

individually adjusting the X and Z' dimensions by hand grinding of each plate or by predetermining the proper dimensions and using mass production methods of precise machine grinding. The advantages of predimensioned crystal units is the insurance of proper operation over a wide temperature range and uniformity of activity. The experience of most manufacturers of low frequency crystal units in the broadcast range and high frequency crystals requiring high activity over a wide temperature range has been that it is necessary to use specific dimensions to insure low rejects in the final tests.

6.53 *AT-Type Crystals*

The modes of motion encountered in the AT-cut crystal are the same as that of the BT-cut. The effects of coupling between most modes is greater

due to the increased piezo electric constant for this particular cut, and the frequency constants are different due to the change in angle with respect to the crystallographic axes. The three series of interfering modes as described for the BT-cut have been measured for this crystal and as shown in Section 6.4 are

$$f_{xf} = \frac{133.84}{X} n_{xf}$$

$$f_{xs} = \frac{254.20}{X} n_{xs} \qquad 6.31$$

$$f_{x's} = \frac{254.00}{Z'} n_{x's}$$

In a manner similar to the BT case a chart has been developed of a folded frequency scale showing the frequencies at which even order $X_{y'}$ flexure modes propagated along X and odd order Z'_x shear modes along X as well as odd order Z'_x shear modes along Z' will interfere with the high frequency $X_{y'}$ shear mode for a crystal 1 centimeter square. This is shown in Fig. 6.26. Its use is the same as that described for the BT case. Insufficient experimental work has been done to indicate the relative shift in the flexure and shear modes along the X axis when they approach each other in frequency. Also, most of the use of square plates and experimental work has been confined to the BT-cut crystals and hence no ok regions are shown for this chart.

APPENDIX B

Equation of elastic and piezoelectric constants for rotation of axes about the X axis. ($s = \sin\theta$; $c = \cos\theta$)

$$c'_{11} = c_{11}$$

$$c'_{22} = c_{11}c^4 + c_{33}s^4 + 2(2c_{44} + c_{13})s^2c^2 + 4c_{14}sc^3$$

$$c'_{33} = c_{11}s^4 + c_{33}c^4 + 2(2c_{44} + c_{13})s^2c^2 - 4c_{14}s^3c$$

$$c'_{44} = c_{44} + (c_{11} + c_{33} - 4c_{44} - 2c_{13})s^2c^2 - 2c_{14}(c^2 - s^2)\,sc$$

$$c'_{55} = c_{44}c^2 + c_{66}s^2 + 2c_{14}sc$$

$$c'_{66} = c_{44}s^2 + c_{66}c^2 - 2c_{14}sc$$

$$c'_{12} = c_{12}c^2 + c_{13}s^2 - 2c_{14}sc$$

$$c'_{13} = c_{12}s^2 + c_{13}c^2 + 2c_{14}sc$$

$$c'_{14} = c_{14}(c^2 - s^2) + (c_{12} - c_{13})sc$$

$$c'_{23} = c_{13}(c^4 + s^4) + (c_{11} + c_{33} - 4c_{44})s^2c^2 - 2c_{14}(c^2 - s^2)sc$$
$$c'_{24} = c_{14}(4s^2 - 1)c^2 + [c_{11}c^2 - c_{33}s^2 - (2c_{44} + c_{13})(c^2 - s^2)]sc$$
$$c'_{34} = -c_{14}(4c^2 - 1)s^2 + [c_{11}s^2 - c_{33}c^2 + (2c_{44} + c_{13})(c^2 - s^2)]sc$$
$$c'_{56} = c_{14}(c^2 - s^2) + (c_{66} - c_{44})sc$$
$$c'_{15} = c'_{16} = c'_{25} = c'_{26} = c'_{35} = c'_{36} = c'_{45} = c'_{46} = 0$$
$$s'_{11} = s_{11}$$
$$s'_{22} = s_{11}c^4 + s_{33}s^4 + (s_{44} + 2s_{13})s^2c^2 + 2s_{14}sc^3$$
$$s'_{33} = s_{11}s^4 + s_{33}c^4 + (s_{44} + 2s_{13})s^2c^2 - 2s_{14}s^3c$$
$$s'_{44} = s_{44} + 4(s_{11} + s_{33} - s_{44} - 2s_{13})s^2c^2 - 4s_{14}(c^2 - s^2)sc$$
$$s'_{55} = s_{44}c^2 + s_{66}s^2 + 4s_{14}sc$$
$$s'_{66} = s_{44}s^2 + s_{66}c^2 - 4s_{14}sc$$
$$s'_{12} = s_{12}c^2 + s_{13}s^2 - s_{14}sc$$
$$s'_{13} = s_{12}s^2 + s_{13}c^2 + s_{14}sc$$
$$s'_{14} = s_{14}(c^2 - s^2) + 2(s_{12} - s_{13})sc$$
$$s'_{23} = s_{13}(c^4 + s^4) + (s_{11} + s_{33} - s_{44})s^2c^2 - s_{14}(c^2 - s^2)sc$$
$$s'_{24} = s_{14}(4s^2 - 1)c^2 + [2(s_{11}c^2 - s_{33}s^2) - (s_{44} + 2s_{13})(c^2 - s^2)]sc$$
$$s'_{34} = -s_{14}(4c^2 - 1)s^2 + [2(s_{11}s^2 - s_{33}c^2) + (s_{44} + 2s_{13})(c^2 - s^2)]sc$$
$$s'_{56} = 2s_{14}(c^2 - s^2) + (s_{66} - s_{44})sc$$
$$s'_{15} = s'_{16} = s'_{25} = s'_{26} = s'_{35} = s'_{36} = s'_{45} = s'_{46} = 0$$
$$d'_{11} = d_{11}$$
$$d'_{12} = -(d_{14}s + d_{11}c)c$$
$$d'_{13} = (d_{14}c - d_{11}s)s$$
$$d'_{14} = d_{14}(c^2 - s^2) - 2d_{11}sc$$
$$d'_{25} = -(d_{14}c + 2d_{11}s)c$$
$$d'_{26} = (d_{14}s - 2d_{11}c)c$$
$$d'_{35} = -(d_{14}c + 2d_{11}s)s$$
$$d'_{36} = (d_{14}s - 2d_{11}c)s$$
$$d'_{15} = d'_{16} = d'_{21} = d'_{22} = d'_{23} = d'_{24} = d'_{31} = d'_{32} = d'_{33} = d'_{34} = 0$$

$$e'_{11} = e_{11}$$

$$e'_{12} = -(2e_{14}s + e_{11}c)c$$

$$e'_{13} = (2e_{14}c - e_{11}s)s$$

$$e'_{14} = e_{14}(c^2 - s^2) - e_{11}sc$$

$$e'_{25} = -(e_{14}c + e_{11}s)c$$

$$e'_{26} = (e_{14}s - e_{11}c)c$$

$$e'_{35} = -(e_{14}c + e_{11}s)s$$

$$e'_{36} = (e_{14}s - e_{11}c)s$$

$$e'_{15} = e'_{16} = e'_{21} = e'_{22} = e'_{23} = e'_{24} = e'_{31} = e'_{32} = e'_{33} = e'_{34} = 0$$

Response of a Linear Rectifier to Signal and Noise*

By W. R. BENNETT

WHEN the input to a rectifier contains both signal and noise components, the resultant output is a complicated non-linear function of signal and noise. Given the spectra of the signal and noise input waves, the law of rectification, and the transmission characteristics of the input and output circuits of the rectifier, it should, in general, be possible to describe the spectrum of the resultant output wave. Before discussing the solution of the general problem, we shall derive some results of a simpler nature, which do not require a consideration of the distribution of the signal and noise energies as functions of frequency.

I. Direct-Current Component of Output

A quantity of considerable importance is the average value of the output amplitude. This is the quantity which would be read by a direct-current meter. Calculation of the average or d-c response can be performed in terms of the distribution of instantaneous output amplitudes in time. The distribution of output amplitude can be computed from the distribution of instantaneous input amplitudes and the law of rectification.

As an example, we shall compute the average current obtained from a linear rectifier when the input to the rectifier consists of a sinusoidal signal with random noise superposed upon it. The probability density function of the signal voltage is first determined, and the result given in (3). The corresponding probability density for the voltage of the noise is well known and is given in (4). The distribution of occurrence of the resultant instantaneous amplitudes of the combined noise and signal voltages is then computed by the rules of mathematical probability, and the result is shown in (7). The assumption that the rectifier is linear then leads directly to an integral which yields the average current obtained from the rectifier.

Let the signal voltage, E_s, be given by

$$E_s = P_o \cos \omega t. \tag{1}$$

The possible angular values of ωt are uniformly distributed throughout the range 0 to 2π. The range E_s to $E_s + dE_s$ corresponds to the range of values of ωt comprised in the interval.

$$\text{arc cos} \frac{E_s}{P_o} < \omega t < \text{arc cos} \frac{E_s + dE_s}{P_o} \tag{2}$$

*Published in *Acous. Soc. Amer. Jour.*, Jan., 1944.

The angular width of this interval is $(P_o^2 - E_s^2)^{-1/2} dE_s$. There are two such intervals in the range $0 < \omega t < 2\pi$. Values of E_s outside the range $-P_o$ to P_o do not exist. Hence, the probability that the signal voltage lies in the interval dE_s at any particular E_s is given by

$$\Phi_s(E_s)dE_s = \left.\begin{cases} 0, & |E_s| > P_o \\ 2(P_o^2 - E_s^2)^{-1/2}\, dE_s/2\pi, & |E_s| < P_o \end{cases}\right\} dE_s \tag{3}$$

Random noise as discussed in this section may be characterized by the fact that the instantaneous amplitudes are normally distributed in time; that is, if $\Phi_n(z)\, dz$ is the probability that the noise amplitude lies in the amplitude interval of width dz at z,

$$\Phi_n(z) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-z^2/2\sigma^2} \tag{4}$$

where σ is the root mean square noise amplitude. The mean noise power dissipated in unit resistance is given by $W_n = \sigma^2$. The corresponding mean signal power is given by $W_s = P_o^2/2$. Let $\Phi_r(z)$ represent the probability density function of the instantaneous sum of the signal and noise amplitudes. Then

$$\Phi_r(z)dz = dz \int_{-\infty}^{\infty} \Phi_s(\lambda)\, \Phi_n(z - \lambda)d\lambda \tag{5}$$

or

$$\Phi_r(z) = \frac{1}{\pi\sigma\sqrt{2\pi}} \int_{P_o}^{P_o} \frac{e^{-(z-\lambda)^2/2\sigma^2}\, d\lambda}{\sqrt{P_o^2 - \lambda^2}} \tag{6}$$

By the substitution $\lambda = P_o \cos\theta$, we may convert the integral to the form

$$\Phi_r(z) = \frac{1}{\pi\sigma\sqrt{2\pi}} \int_o^{\pi} e^{-(z - P_o\cos\theta)^2/2\sigma^2}\, d\theta \tag{7}$$

Suppose we insert a half-wave linear rectifier in series with the source of signal and noise, so that the current I is given in terms of the resultant instantaneous voltage E by

$$I = \begin{cases} 0, & E < 0 \\ \alpha E, & E > 0 \end{cases} \tag{8}$$

Then the average value of current flowing in the circuit is

$$\begin{aligned} \bar{I} &= \alpha \int_0^{\infty} z\Phi_r(z)\, dz \\ &= \frac{\alpha}{\pi\sigma\sqrt{2\pi}} \int_0^{\infty} z\, dz \int_o^{\pi} e^{-(z - P_o\cos\theta)^2/2\sigma^2}\, d\theta \end{aligned} \tag{9}$$

The value of this integral is shown in Appendix I to be

$$\bar{I} = \alpha \sqrt{\frac{W_n}{2\pi}}\, e^{-W_s/2W_n}\left\{I_0(W_s/2W_n) + \frac{W_s}{W_n}\left[I_0\left(\frac{W_s}{2W_n}\right) + I_1\left(\frac{W_s}{2W_n}\right)\right]\right\} \tag{10}$$

This form is particularly convenient for calculation since Watson's Theory of Bessel Functions, Table II, gives $e^{-z}I_0(z)$ and $e^{-z}I_1(z)$ directly.

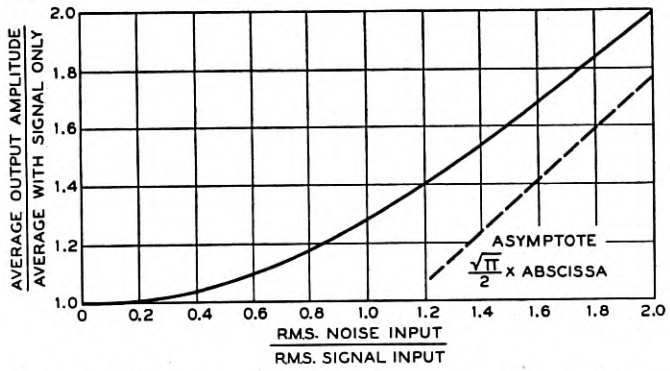

Fig. 1—Variation of direct-current component in response of linear rectifier with ratio of noise input to signal input.

Limiting forms of this equation may be expressed in terms of series in powers of W_s/W_n when the signal power is small compared with the noise power and in powers of W_n/W_s when the noise power is small compared with the signal power. The ascending series for small signal is:

$$\bar{I} = \alpha \sqrt{\frac{W_n}{2\pi}}\left[1 + \frac{1}{2(1!)^2}\frac{W_s}{W_n} + \frac{1(-1)}{2^2(2!)^2}\frac{(W_s)^2}{(W_n)} + \frac{1(-1)(-3)}{2^3(3!)^2}\frac{(W_s)^3}{(W_n)} + \cdots\right] = \alpha \sqrt{\frac{W_n}{2\pi}}\,{}_1F_1\left(\frac{-1}{2}\,;1;\frac{-W_s}{W_n}\right) \tag{11}$$

The asymptotic series, which is available for computation when the signal is large, is

$$\bar{I} \sim \frac{\alpha\sqrt{2W_s}}{\pi}\left[1 + \frac{(-1)^2 W_n}{1!\,4W_s} + \frac{(-1)^2\cdot 1^2}{2!}\frac{(W_n)^2}{(4W_s)} + \frac{(-1)^2\cdot 1^2\cdot 3^2}{3!}\frac{(W_n)^3}{(4W_s)} + \frac{(-1)^2\cdot 1^2\cdot 3^2\cdot 5^2}{4!}\frac{(W_n)^4}{(4W_s)} + \cdots\right] \tag{12}$$

Curves of $\bar{I}$ have been plotted in three ways. Fig. 1 shows the ratio of $\bar{I}$ to $\bar{I}_{so} = \alpha P_o/\pi$, the average current in the absence of noise, as a function

of ratio of rms noise input to rms signal input. Figure 2 shows the ratio of $\bar{I}$ to $\bar{I}_{no} = \alpha\sigma/\sqrt{2\pi}$, the average current in the absence of signal, as a function of ratio of rms signal input to rms noise input. Figure 3 shows

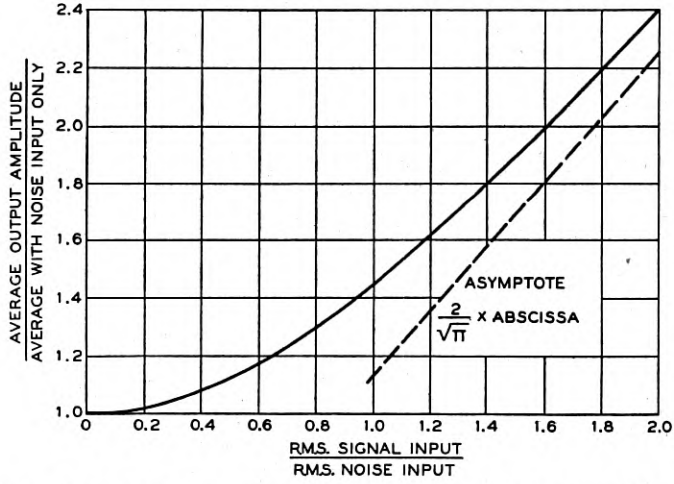

Fig. 2—Variation of direct-current component in response of linear rectifier with ratio of signal input to noise input.

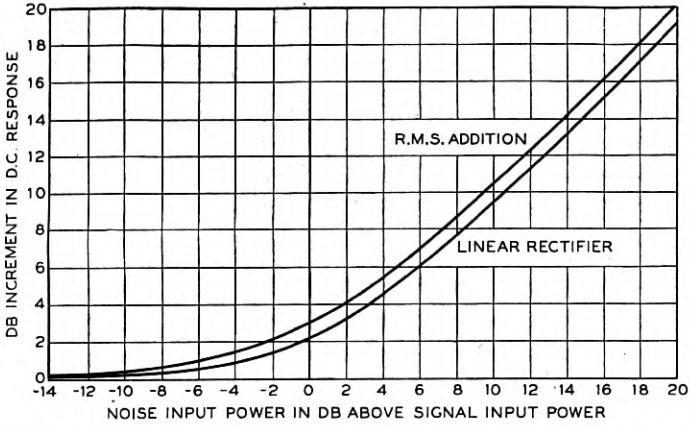

Fig. 3—Variation of direct-current component expressed in decibels, showing comparison between linear rectification and power addition of signal and noise.

the increment in d-c power output in decibels as varying amounts of noise expressed in decibels relative to the signal are added. The corresponding result for power addition is given for comparison.

II. Spectrum of Output

A much more powerful method of attack on this problem is obtained by the use of multiple Fourier series. In this section we shall use Fourier analysis to obtain not only the direct-current output of the rectifier, but also the spectral distribution of the sinusoidal components in the output of the rectifier. We represent the input spectrum by

$$E = P_0 \cos p_0 t + \sum_{n=1}^{N} P_n \cos p_n t \tag{13}$$

This representation is more general than that given by (4) in that a frequency spectrum as well as an amplitude distribution is defined; it may be shown that the probability density for the sum of N sinusoidal waves with incommensurable frequencies approaches (4) when N is large. The first term represents the sinusoidal signal; the mean power which would be dissipated by this signal in unit resistance is

$$W_s = P_0^2/2. \tag{14}$$

The noise is represented by a large number N of sinusoidal components with incommensurable frequencies (or commensurable frequencies with random phase angles) distributed along the frequency range f_1 to f_2 in such a way that the mean noise power in band width Δf is:

$$w(f)\Delta f = \tfrac{1}{2} \sum_{n=\nu(f-f_1)}^{\nu(f+\Delta f-f_1)} P_n^2 = \nu \Delta f \, P^2(f)/2 \tag{15}$$

Here ν is the number of components per unit band width and $P(f)$ represents the amplitude of a component in the neighborhood of frequency f. Note also that the mean total noise input power, W_n, is given by

$$W_n = \int_0^\infty w(f)\, df = \frac{\nu}{2} \int_0^\infty P^2(f)\, df \tag{16}$$

The linear rectifier is specified by the current-voltage relationship (8), which is equivalent to

$$I = -\frac{\alpha}{2\pi} \int_C e^{iEz} \frac{dz}{z^2} \tag{17}$$

where C is an infinite contour going from $-\infty$ to $+\infty$ with an indentation below the pole at the origin. We may expand I in the multiple Fourier series[1]

[1] Bennett and Rice, "Note on Methods of Computing Modulation Products," *Phil. Mag.*, Sept. 1934. The present application represents an extension to N variables of the theory there given for two.

$$I = \sum_{m_0=0}^{\infty} \sum_{m_1=0}^{\infty} \cdots \sum_{m_N=0}^{\infty} a_{m_0 m_1 \cdots m_N} \cos m_0 x_0 \cos m_1 x_1 \cdots \cos m_N x_N \tag{18}$$

where

$$x_k = p_k t, \qquad k = 0, 1, 2, \cdots N \tag{19}$$

$$a_{m_0 m_1 \cdots m_N} = \frac{\epsilon_{m_0} \epsilon_{m_1} \cdots \epsilon_{m_N}}{\pi^{N+1}} \int_0^{\pi} dx_0 \int_0^{\pi} dx_1 \cdots \int_0^{\pi} I \cos m_0 x_0 \cos m_1 x_1 \cdots \cos m_N x_N \, dx_N \tag{20}$$

$$\epsilon_j = \begin{cases} 2, & j \neq 0 \\ 1, & j = 0 \end{cases} \tag{21}$$

The response of the rectifier is thus seen to consist of all orders of modulation products of signal and noise. In a typical case of interest the band of input frequencies is relatively narrow and centered about a high frequency while the output band includes only low frequencies. In such a case the important components in the output are the beats between signal and noise components and between noise components. The *d-c* component is present in the output only if the pass band of the system actually includes zero frequency; we have already computed its value in Section I, but we will derive it again by the method used here as a check.

The amplitude of the d-c component is in fact:

$$a_{00\cdots 0} = -\frac{\alpha}{2\pi} \int_C \frac{J_0(P_0 z) \prod_{n=1}^{N} J_0(P_k z)}{z^2} \, dz, \tag{22}$$

on substitution of the expression for E in the integral representation of I, substituting the result in (20) and interchanging the order of integration. When N is large, P_n is small, hence the principal contribution to the integral occurs near small values of z, where $J_o(P_n z)$ is nearly equal to unity, since the product of a large number of factors, all less than unity, will be small indeed unless each factor is only slightly less than unity. We therefore replace $J_o(P_n z)$ by a function which coincides with it near $z = 0$ and goes rapidly to zero as we depart from this region. Such an approximation (Laplace's process[2]) is

$$J_0(P_n z) \doteq e^{-P_n^2 z^2/4} \tag{23}$$

[2] Watson, "Theory of Bessel Functions," p. 421.

which is correct for the first two terms in the Taylor series expansion near $z = 0$. Therefore, when P_n approaches zero as N approaches infinity,

$$a_{00\ldots0} = \bar{I} = -\frac{\alpha}{2\pi}\int_C J_0(P_0 z)e^{-\sum_{n=1}^{N} P_n^2 z^2/4}\frac{dz}{z^2}$$

$$= -\frac{\alpha}{2\pi}\int_C J_0(P_0 z)e^{-W_n z^2/2}\frac{dz}{z^2} \tag{24}$$

The contour integral cannot be replaced by a real integral directly because the integrand goes to infinity at the origin. However, since

$$\frac{J_0(u)}{u^2} = -\frac{J_1(u)}{u} - \frac{d}{du}\frac{J_0(u)}{u} \tag{25}$$

$$\frac{J_0(Pz)}{z^2} = -\frac{J_1(Pz)}{P^2z^2} - \frac{d}{d(Pz)}\frac{J_0(Pz)}{Pz} = -\frac{J_1(Pz)}{P^2z^2} - \frac{1}{P^2}\frac{d}{dz}\frac{J_0(Pz)}{z} \tag{26}$$

we can substitute (26) in the integral and perform an integration by parts to give the result.

$$\bar{I} = \frac{\alpha}{\pi}\int_0^\infty e^{-W_n z^2/2}\left[\frac{P_0 J_1(P_0 z)}{z} + W_n J_0(P_0 z)\right]dz$$

$$= \alpha\sqrt{\frac{W_n}{2\pi}}\left[{}_1F_1\left(\frac{1}{2};1;-\frac{W_s}{W_n}\right) + \frac{W_s}{W_n}\,{}_1F_1\left(\frac{1}{2};2;-\frac{W_s}{W_n}\right)\right] \tag{27}$$

by Hankel's formula.[3] But it may be shown that (see Appendix II)

$${}_1F_1\left(\tfrac{1}{2};1;-u\right) = e^{-u/2}\,I_0\left(\frac{u}{2}\right) \tag{28}$$

$${}_1F_1\left(\tfrac{1}{2};2;-u\right) = e^{-u/2}\left[I_0(u/2) - I_1(u/2)\right] \tag{29}$$

Hence,

$$\bar{I} = \alpha\sqrt{\frac{W_n}{2\pi}}\,e^{-W_s/2W_n}\left\{I_0(W_s/2W_n) + \frac{W_s}{W_n}\left[I_0(W_s/2W_n) + I_1(W_s/2W_n)\right]\right\} \tag{30}$$

which is identical with the result of Section I, noting that $\sigma = \sqrt{W_n}$. We point out that a resistance-capacity coupled amplifier will not pass this component since there is no transmission at zero frequency.

[3] Watson, "Theory of Bessel Functions," p. 393. As pointed out by Watson, in a footnote, the difficulty with singularities at the origin could be avoided by expressing Hankel's formula in terms of a contour integral instead of an ordinary integral along the real axis. This procedure would lead directly to the hypergeometric function given in (11).

The amplitude of the typical difference product between the signal and the rth noise component is

$$A_{sn} = \tfrac{1}{2} a_{100\cdots010\cdots0}$$
$$= \frac{\alpha}{\pi} \int dz \, \frac{J_1(P_0 z) J_0(P_1 z) J_0(P_2 z) \cdots J_1(P_n z) \cdots J_0(P_N z)}{z^2} \tag{31}$$

Using the same process as before, we replace $J_1(P_n z)$ by

$$J_1(P_n z) \doteq \frac{P_n z}{2} e^{-P_n^2 z^2/8} \tag{32}$$

and obtain in the limit as N becomes indefinitely large

$$A_{sn} = \frac{\alpha P_n}{\pi} \int_0^\infty \frac{J_1(P_0 z)}{z} e^{-W_n z^2/2} \, dz$$
$$= \frac{\alpha P_n}{2} \sqrt{\frac{W_s}{\pi W_n}} \, {}_1F_1\left(\frac{1}{2}; 2; -\frac{W_s}{W_n}\right) \tag{33}$$
$$= \frac{\alpha P_n}{2} \sqrt{\frac{W_s}{\pi W_n}} \, e^{-W_s/2W_n} \left[I_0\left(\frac{W_s}{2W_n}\right) + I_1\left(\frac{W_s}{2W_n}\right) \right]$$

Relations between the ${}_1F_1$ function and Bessel functions are discussed in Appendix II.

The shape of the spectrum of the beats between P_0 and the noise input evidently consists of the superposition of the noise spectra above and below p_o, so that if we write $w_{sn}(f)\,\Delta f$ for the mean energy from this source in that part of the filter output lying in the band of width Δf at f,

$$w_{sn}(f)\Delta f = \frac{\nu \Delta f}{2} \left[(A_{sn}^+)^2 + (A_{sn}^-)^2 \right] \tag{34}$$

$$A_{sn}^+ = [A_{sn}]_{p_n = p_0 + 2\pi f} \tag{35}$$

$$A_{sn}^- = [A_{sn}]_{p_n = p_0 - 2\pi f} \tag{36}$$

$$P_n = \sqrt{\frac{2w(f_n)}{\nu}} \tag{37}$$

$$w_{sn}(f) = \frac{\alpha^2 W_s}{4\pi W_n} e^{-W_s/W_n} \left[I_0\left(\frac{W_s}{2W_n}\right) + I_1\left(\frac{W_s}{2W_n}\right) \right]^2 \times [w(f_0 + f) + w(f_0 - f)] \tag{38}$$

The total noise from this source in the output of a particular filter of transfer admittance $Y(f)$ is obtained by integrating $w_{sn}(f)Y(f)df$ throughout the band of the filter. In the particular case in which the original band of noise is

symmetrical about f_o and occupies the range $f_o - f_a$ to $f_0 + f_a$ and an ideal low pass filter cutting off at $f = f_a$ is used in the rectifier output, the total noise output from beats between signal and noise is

$$W_{sn} = 2\int_0^{f_a} w_{sn}(f)\,df = \frac{\alpha^2 w_s}{4\pi}\, e^{-W_s/W_n}[I_0(W_s/2W_n) + I_1(W_s/2W_n)]^2 \quad (39)$$

Next we shall calculate the spectrum of the energy resulting from beats between individual noise components. We write

$$\begin{aligned} A_{nn} &= \tfrac{1}{2}\, a_{00\cdots010\cdots010\cdots0} \\ &= \frac{\alpha}{\pi}\int_C dz\, \frac{J_0(P_0 z)J_0(P_1 z)\cdots J_1(P_r z)\cdots J_1(P_s z)\cdots J_0(P_N z)}{z^2} \\ &= \frac{\alpha P_r P_s}{2\pi}\int_0^\infty J_0(P_0 z)e^{-W_n z^2/2}\,dz \\ &= \frac{\alpha P_r P_s}{2\sqrt{2\pi W_n}}\,{}_1F_1\left\{\frac{1}{2}; 1; -\frac{W_s}{W_n}\right\} \\ &= \frac{\alpha P_r P_s}{2\sqrt{2\pi W_n}}\, e^{-W_s/2W_n}\, I_0(W_s/2W_n) \end{aligned} \quad (40)$$

To find the resulting spectrum $w_{nn}(f)df$ produced at f by the resultant of all such components, we note that we may sum over all components by beating each component of the primary band with the frequency f above it and adding the resultant power values. The result is

$$w_{nn}(f) = \frac{\alpha^2}{4\pi W_n}\, e^{-W_s/W_n}\, I_0^2(W_s/2W_n)\int_0^\infty w(\lambda)w(\lambda + f)\,d\lambda \quad (41)$$

In the particular case of a flat band of energy extending from f_1 to f_2,

$$\int_0^\infty w(\lambda)w(\lambda + f)\,d\lambda = \int_{f_1}^{f_2 - f} \frac{W_n^2}{(f_2 - f_1)^2}\,d\lambda = \frac{f_2 - f_1 - f}{(f_2 - f_1)^2}\, W_n^2, \quad (42)$$

$$0 < f < f_2 - f_1$$

$$w_{nn}(f) = \frac{\alpha^2(f_2 - f_1 - f)W_n}{4\pi\,(f_2 - f_1)^2}\, e^{-W_s/W_n}\, I_0^2(W_s/2W_n), \quad (43)$$

$$0 < f < f_2 - f_1$$

The total mean power of this type lying in the band 0 to f_b is

$$W_{nn}(f_b) = \int_0^{f_b} w_{nn}(f)\,df = \frac{\alpha^2 W_n(f_2 - f_1 - f_b/2)f_b}{4\pi(f_2 - f_1)^2}\, e^{-W_s/W_n} \quad (44)$$

$$I_0^2(W_s/2W_n)$$

provided $f_b < f_2 - f_1$. The spectrum is confined to the region $0 < f < f_2 - f_1$. If f_b is equal to $f_2 - f_1$ so that the output filter passes all the noise of this type, we have

$$W_{nn}(f_2 - f_1) = W_{nn} = \frac{\alpha^2 W_n}{8\pi} e^{-W_s/W_n} I_0^2(W_s/2W_n) \tag{45}$$

This result seems to hold approximately for a considerable range of input spectra. For example, if we assume that the original noise is shaped like an error function about f_o, i.e.,

$$w_n(f) = W_n \sqrt{a/\pi}\, e^{-a(f-f_0)^2} \tag{46}$$

with f taken from $-\infty$ to $+\infty$ with small error for large f_o,

$$\int_{-\infty}^{\infty} w(\lambda) w(\lambda + f)\, d\lambda = W_n^2 \sqrt{a/2\pi}\, e^{-af^2/2} \tag{47}$$

$$\int_0^{\infty} df \int_{-\infty}^{\infty} w(\lambda) w(\lambda + f)\, d\lambda = W_n^2/2 \tag{48}$$

which is in agreement with (45).

The output of a half-wave linear rectifier contains fundamental components and all even order modulation products. In general, the amplitudes of the higher order products are small compared with the lower order. In a particular problem some consideration of where the principal products fall in the frequency band is required. The products just considered give a fair approximation for the problem of detection of a radio frequency band of signal and noise followed by audio amplification. Contain other products should also be added to obtain higher accuracy. We have calculated the products of order zero and two; the next ones of importance are the fourth order, since the third order products vanish in a perfectly linear rectifier. The fourth order products in this case which fall in the audio band are of frequency $2p_o - p_r - p_s$, $p_o + p_q - p_r - p_s$, and $p_n + p_q - p_r - p_s$, where the subscripts n, q, r, s refer to the original noise component frequencies. The latter is, however, less important than the sixth order product $3p_o - p_q - p_r - p_s$, which involves only three noise components, Expressions for the contributions from these products are given in Appendix III.

Figure 4 shows computed curves for the noise produced in an audio band by the various components. Curve A is $W_{sn} + W_{nn}$ and includes what are usually regarded as the principal contributors, the difference frequencies between signal and noise, and between individual noise components. Curve B is obtained by adding to Curve A, the contribution from the fourth order products $2p_o - p_r - p_s$ and $p_o + p_q - p_r - p_s$ and the sixth order products $3p_o - p_q - p_q - p_s$. Thus all products which include three or less noise fundamental components are included. The curves are plotted in terms of

fraction of noise power received compared to the limiting noise when the mean signal input power is made indefinitely large compared to the mean input noise power. Some experimental points given by Williams[4] are shown for comparison. Williams gives the intercept at zero signal power as 35%; the theoretical value deduced here is $\pi/8$ or 39.27%. It will be noted that the inclusion of the higher order products improves the agreement between experimental and theoretical curves, even though the value of the intercept is unaffected by them. It shold also be stressed that our analysis applies

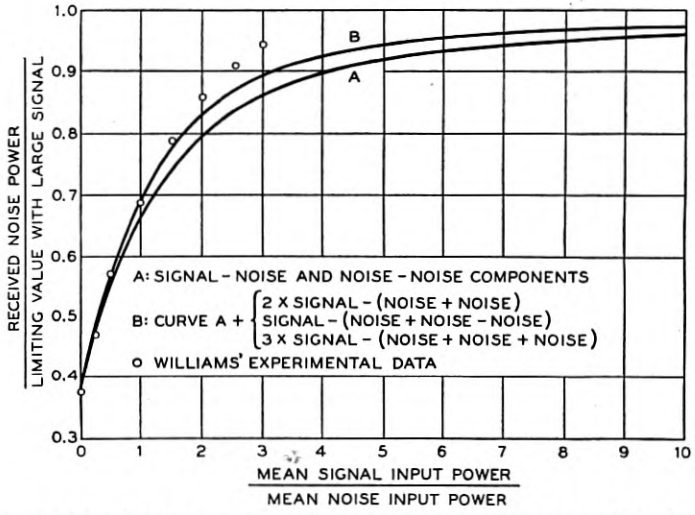

Fig. 4—Calculated noise power in audio band of output of linear rectifier when noise and signal are applied in a relatively narrow high-frequency band. The direct-current component is excluded.

strictly to purely resistive networks. The conventional radio detector circuit (which was used by Williams), in which a condenser is shunted across a resistance in series with a diode, departs from the conditions here assumed because of the reactive element, the condenser. The customary approximation made in treating this circuit is that the condenser has infinite impedance in the audio frequency range and zero impedance at the radio frequencies. This leads to a bias on the detector which depends on the signal. The methods given here may be applied, but the resulting formulas are much more difficult from the standpoint of numerical computation.

A recent paper by Ragazzini[5] gives an approximate solution based on

[4] F. C. Williams, "The Response of Rectifiers to Fluctuation Voltages," *Journal I. E. E.*, 1937, Vol. 80, pp. 218–226.

[5] John Ragazzini, "The Effect of Fluctuation Voltages on the Linear Detector," *Proc. I. R. E.*, June 1942, Vol. 30, p. 277–288.

expanding the envelope of the input wave by the binomial theorem and retaining only the first two terms. The validity depends on the noise amplitude being small compared with the sum of signal and noise, and hence the result should agree with our solution in the neighborhood of $W_n/W_s = 0$, which it does. When W_s/W_n is small, the error is appreciable. Ragazzini's result (Equation 15 of the paper) expressed in our notation is

$$W_{sn} + W_{nn} \doteq \frac{\alpha^2}{\pi^2} \frac{W_n(1 + \frac{1}{2} W_n/W_s)}{1 + W_n/W_s} \tag{49}$$

It will be seen by comparing the limiting values for $W_s/W_n = 0$ with that of $W_s/W_n = \infty$ from (49) that the intercept of the curve of Fig. 4 would be 50% instead of our value of 39.27%.

The results given in the present paper have been compiled from unpublished memoranda and notes by the author extending back as far as 1935. Discussions with colleagues have been of great aid, and in particular acknowledgment is made to Messrs. S. O. Rice and R. Clark Jones for many helpful suggestions.

APPENDIX I

EVALUATION OF INTEGRAL FOR I

Interchanging the order of integration in (9), we have

$$\bar{I} = \frac{\alpha}{\pi\sqrt{2\pi W_n}} \int_0^\pi d\theta \int_0^\infty e^{-(z - P_0 \cos\theta)^2/2W_n} z\, dz \tag{50}$$

By substituting $z = P_0 \cos\theta + u\sqrt{2W_n}$, we may evaluate the second integral in terms of the error function, obtaining

$$\bar{I} = \frac{\alpha}{\pi^{3/2}} \int_0^\pi d\theta \int_{-P\cos\theta/\sqrt{2W_n}}^\infty e^{-u^2} (u\sqrt{2W_n} + P_0 \cos\theta)\, du$$

$$= \frac{\alpha}{\pi} \frac{\sqrt{W_n}}{2\pi} \int_0^\pi e^{-P_0^2 \cos^2\theta/2W_n}\, d\theta$$

$$+ \frac{\alpha P_0}{2\pi} \int_0^\pi \operatorname{erf}(P_0 \cos\theta/\sqrt{2W_n}) \cos\theta\, d\theta$$

$$= \frac{\alpha}{\pi} \frac{\sqrt{W_n}}{2\pi} e^{-P_0^2/4W_n} \int_0^\pi e^{-\cos 2\theta/4W_n}\, d\theta$$

$$+\frac{\alpha P_0}{2\pi}\int_0^\pi \frac{d}{d\theta}\left[\operatorname{erf}\left(\frac{P_0\cos\theta}{\sqrt{2W_n}}\right)\sin\theta\right]d\theta - \frac{\alpha P_0}{2\pi}\int_0^\pi \sin\theta\,\frac{d}{d\theta}\left(\operatorname{erf}\frac{P_0\cos\theta}{\sqrt{2W_n}}\right)d\theta$$

$$= \frac{\alpha}{2\pi}\frac{\sqrt{W_n}}{2\pi}e^{-W_s/2W_n}\int_0^{2\pi} e^{-W_s\cos\Phi/2W_n}\,d\Phi$$

$$+\frac{\alpha W_s}{2\pi\sqrt{2\pi W_n}}\int_0^{2\pi} e^{-W_s\cos\Phi/2W_n}(1-\cos\Phi)\,d\Phi$$

$$= \alpha\sqrt{\frac{W_n}{2\pi}}\,e^{-W_s/2W_n}\left(I_0(W_s/2W_n)\right.$$

$$\left.+\frac{W_s}{W_n}\left[I_0(W_s/2W_n)+I_1(W_s/2W_n)\right]\right) \tag{10}$$

In the above we have made use of the relations:

$$\operatorname{erf} z = \frac{2}{\sqrt{\pi}}\int_0^z e^{-z^2}\,dz \tag{51}$$

$$\frac{d}{dz}\operatorname{erf} z = \frac{2}{\sqrt{\pi}}e^{-z^2} \tag{52}$$

$$\int_0^{2\pi} e^{-z\cos\Phi}\cos m\Phi\,d\Phi = (-)^m 2\pi I_m(z) \tag{53}$$

APPENDIX II

RELATIONS BETWEEN HYPERGEOMETRIC AND BESSEL FUNCTIONS

The modulation coefficients appearing in the linear rectification of noise are expressible in compact form in terms of the hypergeometric function:

$$\begin{aligned} {}_1F_1(a;c;-z) &= 1 - \frac{a}{c}\frac{z}{1!} + \frac{a(a+1)}{c(c+1)}\frac{z^2}{2!} - \cdots \\ &= \frac{\Gamma(c)}{\Gamma(a)}\sum_{m=0}^{\infty}\frac{\Gamma(a+m)}{\Gamma(c+m)m!}(-z)^m \end{aligned} \tag{54}$$

The ${}_1F_1$ function is a limiting case of the more familiar Gaussian hypergeometric function ${}_2F_1(a, b; c; z)$, viz.

$${}_1F_1(a;c,z) = \operatorname*{Limit}_{b=\infty} {}_2F_1(a, b; c; z/b) \tag{55}$$

In certain special cases this function may be expressed in terms of exponential and Bessel functions. For example, by a formula given by

Campbell and Foster, Fourier Integrals for Practical Application, Bell System Monograph B-584, p. 32 (also Watson, Theory of Bessel Functions, p. 191), we may show that

$$ {}_1F_1\left(\nu + \frac{1}{2}; 2\nu + 1; -z\right) = \frac{2^{2\nu}\Gamma(\nu + 1)e^{-z/2}}{(-z)^{\nu}} I_{\nu}(-z/2) \tag{56} $$

or setting $\nu = 0$

$$ {}_1F_1(\tfrac{1}{2}; 1; -z) = e^{-z/2} I_0(z/2) \tag{57} $$

which is one of the functions appearing in our work.

We have also encountered the function ${}_1F_1(1/2; 2; -z)$ which is not directly reducible by the above formula. The reduction may be effected in a number of ways. By making use of the relation obtained from (56) by setting $\nu = 1$,

$$ {}_1F_1(3/2; 3; -z) = \frac{4}{z} e^{-z/2} I_1(z/2) \tag{58} $$

and noting that

$$ \begin{aligned} &{}_1F_1(1/2; 2; -z) - {}_1F_1(1/2; 1; -z) \\ &= \frac{1}{\Gamma(1/2)} \sum_{m=0}^{\infty} \frac{\Gamma(m + 1/2)}{m!(m + 1)!} (-z)^m - \frac{1}{\Gamma(1/2)} \sum_{m=0}^{\infty} \frac{\Gamma(m + 1/2)}{(m!)^2} (-z)^m \\ &= \frac{-1}{\Gamma(1/2)} \sum_{m=0}^{\infty} \frac{\Gamma(m + 1/2)m}{m!(m + 1)!} (-z)^m \\ &= \frac{z}{\Gamma(1/2)} \sum_{m=0}^{\infty} \frac{\Gamma(m + 3/2)}{(m + 2)!m!} (-z)^m \\ &= \frac{z}{4} {}_1F_1(3/2; 3; -z), \end{aligned} \tag{59} $$

we find that[6]

$$ {}_1F_1(1/2; 2; -z) = e^{-z/2} [I_0(z/2) + I_1(z/2)] \tag{60} $$

It may also be verified by integrating the series directly that

$$ \int_0^z {}_1F_1(1/2; 1; -z)\, dz = z\, {}_1F_1(1/2; 2; -z) \tag{61} $$

Combining this relation with (57) and (60) above, we deduce the indefinite integrals

[6] The relation (60) was brought to the attention of the author by Mr. R. M. Foster.

$$\left.\begin{aligned}
\int e^x I_0(x)\,dx &= xe^x[I_0(x) - I_1(x)] \\
\int e^{-x} I_0(x)\,dx &= xe^{-x}[I_0(x) + I_1(x)] \\
\int e^x I_1(x)\,dx &= e^x[(1-x)I_0(x) + xI_1(x)] \\
\int e^{-x} I_1(x)\,dx &= e^{-x}[(1+x)I_0(x) + xI_1(x)]
\end{aligned}\right\} \tag{62}$$

These integrals may be derived by differentiating the right hand members, and could, therefore, serve as a basis for an alternate derivation of (60).

In addition it was noted in Eq. (11) that the constant term in the modulation spectrum could be expressed in terms of ${}_1F_1\,(-1/2;\ 1;\ -z)$; from the equations given, it follows that we must have the relation:

$${}_1F_1\,(-1/2; 1; -z) = e^{-z/2}\,[(1+z)\,I_0(z/2) + z\,I_1(z/2)] \tag{63}$$

Another interesting set of formulas which can be obtained as a by-product from (62) by setting $x = iy$ is:

$$\left.\begin{aligned}
\int J_0(y)\cos y\,dy &= y[J_0(y)\cos y + J_1(y)\sin y] \\
\int J_0(y)\sin y\,dy &= y[J_0(y)\sin y - J_1(y)\cos y] \\
\int J_1(y)\cos y\,dy &= yJ_1(y)\cos y - J_0(y)(y\sin y - \cos y) \\
\int J_1(y)\sin y\,dy &= yJ_1(y)\sin y + J_0(y)(y\cos y - \sin y)
\end{aligned}\right\} \tag{64}$$

The hypergeometric notation is particularly convenient in determining series expansions for the coefficients to be used for calculation when the variable z is either very small or very large. For small values of z, the form (54) suffices; for large values of z, we may use the general asymptotic expansion formula[7] for the real part of z positive:

$$\begin{aligned}
{}_1F_1(a;\ c;\ -z) &= \frac{\Gamma(c)}{\Gamma(c-a)z^a}\,{}_2F_0(a,\ 1+a-c;\ 1/z) \\
&= \frac{\Gamma(c-a)z^a}{\Gamma(c)}\left[1 + \frac{a(1+a-c)}{1!z}\right. \\
&\qquad \left. + \frac{a(a+1)(1+a-c)(2+a-c)}{2!z^2} + \cdots\right]
\end{aligned} \tag{65}$$

[7] Copson, "Functions of a Complex Variable," pp. 264–5.

The series expansions required here could also be obtained from the appropriate series for Bessel functions. It will be noted, however, that the typical modulation coefficient can be expressed in terms of either a single ${}_1F_1$ function or several Bessel functions, so that manipulations must be performed on the series for the latter to give the final result. The Bessel functions on the other hand are more convenient for numerical computations because of the excellent tables available.

Reduction formulas for certain other hypergeometric functions are needed in evaluating the higher order products. They are:

$$ {}_1F_1(3/2;\,1;\,-z) = e^{-z/2}[(1-z)I_0(z/2) + I_1(z/2)] \tag{66} $$

$$ {}_1F_1(3/2;\,2;\,-z) = e^{-z/2}[I_0(z/2) - I_1(z/2)] \tag{67} $$

$$ {}_1F_1(5/2;\,4;\,-z) = \frac{4}{z}\,e^{-z/2}\left[\left(\frac{4}{z}+1\right)I_1(z/2) - I_0(z/2)\right] \tag{68} $$

Derivation of these is facilitated by the use of the easily demonstrated relations:

$$ {}_1F_1(a;\,1;\,-z) = \frac{d}{dz}\,[z\,{}_1F_1(a;\,2;\,-z)] \tag{69} $$

$$ 2z\,{}_1F_1(a;\,2;\,-z) = \frac{d}{dz}\,[z^2\,{}_1F_1(a;\,3;\,-z)] \tag{70} $$

$$ {}_1F_1(3/2;\,3;\,-z) - {}_1F_1(3/2;\,2;\,-z) = \frac{z}{4}\,{}_1F_1(5/2;\,4;\,-z) \tag{71} $$

APPENDIX III

HIGHER ORDER PRODUCTS

The methods described in Section II may be applied to calculate the general expression for the general modulation coefficient. The result is for the amplitude of the term $\cos mp_0 t \cos p_{n_1} t \cos p_{n_2} t \cdots \cos p_{n_M} t$:

$$ a_{mM} = \frac{(-)^{\frac{m+M}{2}+1}\,P_{n_1}P_{n_2}\cdots P_{n_M}}{\pi(W_n/2)^{(M-1)/2}\,m!}\,\Gamma\left(\frac{m+M-1}{2}\right)\frac{(W_s)^{m/2}}{(W_n)} \times {}_1F_1\left(\frac{m+M-1}{2};\,m+1;\,\frac{-W_s}{W_n}\right) \tag{72} $$

The coefficient of the term $\cos(mp_0 \pm p_{n_1} \pm p_{n_2} \pm \cdots p_{n_M})\,t$ is a_{mM} divided by $2^{M-1}\,\epsilon_m$. The number of terms of a particular type falling in a particular frequency interval can be calculated by a method previously described by

the author.[8] Under the assumed conditions that the original noise spectrum is either flat throughout a limited range, or falls off like an error function, and that the audio amplifier passes all the difference components in question, we find the following results:

$$2p_0 - p_r - p_s:$$

$$W_{2s,nn} = \frac{\alpha^2 W_n}{8\pi} e^{-W_s/W_n} I_1^2(W_s/2W_n) \tag{73}$$

$$p_0 + p_q - p_r - p_s:$$

$$W_{sn,nn} = \frac{\alpha^2 W_s}{32\pi} e^{-W_s/W_n} [I_0(W_s/2W_n) - I_1(W_s/2W_n)]^2 \tag{74}$$

$$3p_0 - p_q - p_r - p_s:$$

$$W_{3s,nnn} = \frac{\alpha^2 W_s}{32\pi} e^{-W_s/W_n} [(1 + 4W_n/W_s)I_1(W_s/2W_n) - I_0(W_s/2W_n)]^2 \tag{75}$$

This includes all beats containing not more than three noise fundamentals. The reductions of hypergeometric functions to exponential and Bessel functions given in Appendix II have been used in deriving the above results.

[8] Bennett, "Cross-Modulation in Multichannel Amplifiers," *Bell Sys. Tech. Jour.*, Oct. 1940, Vol. XIX, pp. 587–610.

Dielectric Constants and Power Factors at Centimeter Wave-Lengths

By CARL R. ENGLUND

The theory underlying the measurement of dielectric constants and power factors, by means of resonant lengths of coaxial transmission line, is developed, apparatus used for such measurements is illustrated and the measurement routine described. A table of typical results is appended together with an "X tan X" table for aid in the calculations.

Introduction

THERE are two instrumentalities available for measuring dielectric constants and power factors at centimeter wave-lengths. These are, coaxial conductor lines and wave guides. Which one is, for any condition, the more favorable one depends a great deal upon the wave-lengths used. Under the conditions encountered in this work the coaxial line appeared to have the practical superiority, down to something like 10 cms. wave-length, anyway. Below this, the wave guide is very manageable and has several advantageous features.

When this work was begun, the most easily available and practicable vacuum tube which would oscillate around 20 cms. wave-length was the W. E. Co. 368A. This could be pushed down to something below 19 cms. wave-length but was undependable there and as a practical compromise 22.5 cms. wave-length was finally chosen. Later another tube became available and as it could be operated down to at least 9 cms. it was used in the more recent work. Thus, while the bulk of the measurements made were at 22.5 cms. wave-length, a good share of the samples investigated were also measured at approximately 10 cms. wave-length.

Any measurements made at these wave-lengths must be made in the form of transmission line measurements and the dielectric must be physically part of the coaxial line. There are various transmission line quantities definable and measurable, such as series impedance per unit length, shunt admittance per unit length, surge impedance, impedance transformation factor, voltage and current step-up factors, resonance selectivity or "Q", etc. The first two are measurable directly only at long wave-lengths, the last two are properties of space resonant line elements. Of these the "Q" was the most advantageous in the present instance.

"Q" Definition

At low frequencies the resonance selectivity factor of lumped circuits is identified as the "Q" and is defined as $\frac{\omega L}{R}$. It is measured by a detuning process. For a length of transmission line with negligible shunt conductance losses this process gives $\frac{\omega L}{R}$ as for a coil; when this process is applied to complex circuits the physical embodiment of the "Q" becomes difficult to realize and it is preferable to define the "Q" in terms of the detuning process itself. This is equally true for the resonant, centimeter wave, line element and we proceed as follows: For this element some current or voltage amplitude, conveniently measurable, is selected and three values of it are measured as the line tuning is varied. This variation may be either in generator frequency for constant line length or in line length for constant generator frequency.

Thus, for example,

$$\left.\begin{aligned} Q &= \frac{f_0}{f_2 - f_1}, \quad \text{where } f_2 > f_0 > f_1 \\ Q &= \frac{\ell_0}{\ell_2 - \ell_1}, \quad \text{where } \ell_2 > \ell_0 > \ell_1 \end{aligned}\right\} A_2^2 = A_1^2 = \frac{A_0^2}{2} \qquad (1)$$

with A_0 as the resonant amplitude. For low-loss lines the two definitions will give the same results in practice. Neither is ideal for second order accuracy since there is a variation of line constants with frequency in the first and a variation in total attenuation in the second.

For practical reasons it is usually preferable to excite and observe the line resonance in terms of the current at one end, this end shorted. The elementary line lengths are then the quarter and the half-wave ones, the former with open circuit far end, the latter with shorted far end. The latter is the more nearly ideal unit. In order to short effectively the input end, the input and output couplings must be made as loose as possible. As these couplings are reduced the observed "Q" will asymptotically approach the line "Q". At the present moment the line variation in length is the most convenient process, the chief trouble being the micrometric measurement of the tiny length changes involved. Thus for 10 cms wave-length and a half-wave coaxial line, a "Q" of 1000 involves a plunger movement of .0019 inches.

Theory of Measurement

It is shown in the appendix that the "Q" of a given resonant line segment can be broken up into parts representing the equivalent "Q's" of the terminal impedances and the line itself. Thus

$$\frac{1}{Q} = \frac{1}{Q_a} + \frac{1}{Q_0} + \frac{1}{Q_\ell} \tag{2}$$

where "Q" is the actually measured quantity, Q_a is the part due to the line itself, Q_0 and Q_ℓ the parts due to the near and far end terminations, respectively.

If we now take a quarter-wave line segment, with near end shorted through a movable plunger and far end open, we may make two "Q" measurements without and with the far end loaded with a dielectric segment, and obtain

$$\left.\begin{aligned} &\left\{\begin{aligned} &\frac{1}{Q'} = \frac{1}{Q_a} + \frac{1}{Q_0} = \frac{d'}{\lambda/4} \\ &\frac{1}{Q} = \frac{1}{Q_a} + \frac{1}{Q_0} + \frac{1}{Q_\ell} = \frac{d}{\lambda/4} \end{aligned}\right\} \\ &\text{and} \quad \frac{1}{Q} - \frac{1}{Q'} = \frac{1}{Q_\ell} = \frac{d - d'}{\lambda/4} \end{aligned}\right\} \left\{\begin{aligned} d' &= \ell_2' - \ell_1' \\ d &= \ell_2 - \ell_1 \end{aligned}\right. \tag{3}$$

with d' and d equal to the widths of the resonance curves halfway down in power. These two d's are, of course, directly measurable.

When the line is loaded with a dielectric segment the loaded part of the line can be represented as an impedance Z_ℓ connected to the unloaded remainder of the line. The effect of the loaded segment upon the unloaded line (See appendix, eq. 4) appears in the form $\frac{\sqrt{\frac{Z}{D}}}{Z_\ell}$ where $\sqrt{\frac{Z}{D}}$ is the surge impedance of the unloaded line, with "Z" and "D" the series impedance and shunt admittance, respectively, for unit length of this line. If we put

$$\tanh \theta = \tanh (a_\ell + ib_\ell) = \frac{\sqrt{\frac{Z}{D}}}{Z_\ell} \tag{4}$$

we have

$$\left\{\begin{aligned} &\frac{1}{Q_\ell} = \frac{d - d'}{\lambda/4} = \frac{4a_\ell}{\pi} \\ &\Delta\ell + t = \frac{\lambda}{2\pi} b_\ell \end{aligned}\right. \tag{5}$$

where $\Delta\ell$ is the measured plunger movement necessary to retune the line, after adding the dielectric loading, and "t" is the length of the dielectric segment.

Now, the power factor of "Z_ℓ" is the same as that of $\frac{\sqrt{\frac{Z}{D}}}{Z_\ell}$, as long as $\sqrt{\frac{Z}{D}}$ is substantially a resistance, and since

$$\tanh (a_\ell + i b_\ell) = \frac{\sinh 2a_\ell + i \sin 2b_\ell}{\cosh 2a_\ell + \cos 2b_\ell}, \tag{6}$$

we have

$$\text{power factor } Z_\ell = \text{p.f.} = \frac{\sinh 2a_\ell}{\sin 2b_\ell}. \tag{7}$$

Substituting eq. (5) in (7),

$$\text{p.f.} = \frac{\sinh \frac{2\pi}{\lambda}(d - d')}{\sin \frac{4\pi}{\lambda}(\Delta\ell + t)}, \tag{8}$$

which is the power factor of the loaded line segment in terms only of measurable lengths.

This does not complete the theory, however. We are interested in the power factor of the dielectric itself and it is evident that except for very short dielectric segments, the variation of the standing electrical field along the dielectric segment will result in a calculated power factor smaller than the true one. We also wish to determine the dielectric constant.

The impedance of the dielectric line segment, open circuited at the far end, can be written as

$$Z_\ell = \frac{\sqrt{\frac{L}{\epsilon C}}}{\tanh \left(\alpha + i\frac{2\pi\sqrt{\epsilon}}{\lambda}\right) t} \tag{9}$$

where "α" is the attenuation per unit length and "ϵ" is the dielectric constant. Hence $\tanh (a_\ell + ib_\ell) = \sqrt{\epsilon} \tanh \left(\alpha + i\frac{2\pi\sqrt{\epsilon}}{\lambda}\right) t$ and

$$\text{p.f.} = \frac{\sinh 2\alpha t}{\sin \frac{4\pi\sqrt{\epsilon}}{\lambda} t} \tag{10}$$

an alternative expression. Now when "t" is very small the functions of the angles become equal to the angles and we write, for the dielectric power factor itself

$$\text{P.F.} = \frac{2\alpha t}{\frac{4\pi\sqrt{\epsilon}}{\lambda} t}. \tag{11}$$

Dividing this expression by eq. (10)

$$\text{P.F.} = \text{p.f.} \frac{\sin \frac{4\pi\sqrt{\epsilon}\, t}{\lambda}}{\frac{4\pi\sqrt{\epsilon}\, t}{\lambda}} \cdot \frac{2\alpha t}{\sinh 2\alpha t}$$

and as the last term is always very nearly unity we have, if we put $\frac{4\pi\sqrt{\epsilon}\, t}{\lambda} = 4X$,

$$\text{P.F.} = \frac{\sinh \frac{2\pi}{\lambda}(d - d')}{\sin \frac{4\pi}{\lambda}(\Delta\ell + t)} \cdot \frac{\sin 4X}{4X}. \tag{12}$$

Ordinarily the "sinh" is very closely equal to the angle.

The reactance of the dielectric segment of line is necessarily equal to the reactance of the part of the original line which it displaces, since space resonance occurs in both cases. Hence,

$$\tan \pi \frac{\Delta\ell + t}{\lambda} = \sqrt{\epsilon} \tan \pi \frac{\sqrt{\epsilon}\, t}{\lambda} \tag{13}$$

which we can rewrite to

$$\frac{\pi t}{\lambda} \cdot \tan \pi \frac{\Delta\ell + t}{\lambda} = \pi \frac{\sqrt{\epsilon}\, t}{\lambda} \cdot \tan \pi \frac{\sqrt{\epsilon}\, t}{\lambda}.$$

Putting

$$\begin{cases} y = \dfrac{\pi t}{\lambda} \tan \pi \dfrac{\Delta\ell + t}{\lambda} \\ X = \dfrac{\pi\sqrt{\epsilon}\, t}{\lambda} \end{cases} \qquad \text{we have} \quad y = X \tan X. \tag{14}$$

"y" is directly determinable by measurement and this gives X from the X tan X table supplied.[1] The value of $\epsilon = \left[\frac{X}{\frac{\pi t}{\lambda}}\right]^2$ follows and P.F. is immediately calculable. This completes the reduction of the observation.

[1] As no X tan X table to the necessary subdivision was available, one was calculated from the Hayashi tan X tables.

X TAN X

X	0	1	2	3	4	5	6	7	8	9
.00	.0000 0000	.0000 0100	.0000 0400	.0000 0900	.0000 1600	.0000 2500	.0000 3600	.0000 4900	.0000 6400	.0000 8100
.01	.0001 0000	.0001 2100	.0001 4401	.0001 6901	.0001 9601	.0002 2502	.0002 5602	.0002 8903	.0003 2403	.0003 6104
.02	.0004 0005	.0004 4106	.0004 8408	.0005 2909	.0005 7611	.0006 2513	.0006 7615	.0007 2918	.0007 8420	.0008 4124
.03	.0009 0027	.0009 6131	.0010 2435	.0010 8940	.0011 5645	.0012 2550	.0012 9656	.0013 6963	.0014 4470	.0015 2177
.04	.0016 0085	.0016 8194	.0017 6504	.0018 5014	.0019 3725	.0020 2637	.0021 1749	.0022 1063	.0023 0577	.0024 0292
.05	.0025 0209	.0026 0326	.0027 0644	.0028 1163	.0029 1884	.0030 2806	.0031 3928	.0032 5252	.0033 6778	.0034 8504
.06	.0036 0433	.0037 2562	.0038 4893	.0039 7426	.0041 0160	.0042 3096	.0043 6234	.0044 9573	.0046 3114	.0047 6857
.07	.0049 0802	.0050 4949	.0051 9298	.0053 3849	.0054 8602	.0056 3557	.0057 8715	.0059 4075	.0060 9637	.0062 5402
.08	.0064 1369	.0065 7539	.0067 3911	.0069 0486	.0070 7264	.0072 4245	.0074 1429	.0075 8815	.0077 6405	.0079 4198
.09	.0081 2194	.0083 0393	.0084 8796	.0086 7402	.0088 6212	.0090 5225	.0092 4442	.0094 3862	.0096 3486	.0098 3315
.10	.0100 3347	.0102 3583	.0104 4023	.0106 4668	.0108 5516	.0110 6570	.0112 7827	.0114 9289	.0117 0956	.0119 2828
.11	.0121 4904	.0123 7185	.0125 9672	.0128 2363	.0130 5259	.0132 8361	.0135 1668	.0137 5181	.0139 8899	.0142 2823
.12	.0144 6952	.0147 1287	.0149 5829	.0152 0576	.0154 5529	.0157 0689	.0159 6055	.0162 1628	.0164 7407	.0167 3393
.13	.0169 9585	.0172 5984	.0175 2591	.0177 9404	.0180 6425	.0183 3653	.0186 1088	.0188 8731	.0191 6582	.0194 4640
.14	.0197 2907	.0200 1381	.0203 0063	.0205 8954	.0208 8053	.0211 7360	.0214 6876	.0217 6601	.0220 6534	.0223 6677
.15	.0226 7028	.0229 7589	.0232 8359	.0235 9339	.0239 0528	.0242 1927	.0245 3535	.0248 5354	.0251 7383	.0254 9619
.16	.0258 2071	.0261 4731	.0264 7602	.0268 0683	.0271 3975	.0274 7479	.0278 1193	.0281 5119	.0284 9256	.0288 3605
.17	.0291 8166	.0295 2939	.0298 7923	.0302 3120	.0305 8529	.0309 4151	.0312 9985	.0316 6032	.0320 2292	.0323 8765
.18	.0327 5452	.0331 2351	.0334 9464	.0338 6791	.0342 4332	.0346 2087	.0350 0056	.0353 8239	.0357 6637	.0361 5250
.19	.0365 4077	.0369 3119	.0373 2377	.0377 1849	.0381 1537	.0385 1441	.0389 1561	.0393 1896	.0397 2448	.0401 3216
.20	.0405 4201	.0409 5402	.0413 6820	.0417 8455	.0422 0307	.0426 2377	.0430 4664	.0434 7169	.0438 9892	.0443 2832
.21	.0447 5991	.0451 9369	.0456 2965	.0460 6780	.0465 0814	.0469 5067	.0473 9540	.0478 4232	.0482 9144	.0487 4275
.22	.0491 9627	.0496 5200	.0501 0992	.0505 7006	.0510 3240	.0514 9696	.0519 6373	.0524 3271	.0529 0391	.0533 7733
.23	.0538 5297	.0543 3084	.0548 1093	.0552 9325	.0557 7779	.0562 6457	.0567 5358	.0572 4483	.0577 3831	.0582 3404
.24	.0587 3201	.0592 3222	.0597 3468	.0602 3939	.0607 4634	.0612 5555	.0617 6702	.0622 8074	.0627 9673	.0633 1497
.25	.0638 3548	.0643 5825	.0648 8330	.0654 1061	.0659 4020	.0664 7207	.0670 0621	.0675 4263	.0680 8133	.0686 2232
.26	.0691 6560	.0697 1117	.0702 5903	.0708 0918	.0713 6163	.0719 1638	.0724 7343	.0730 3278	.0735 9444	.0741 5842
.27	.0747 2470	.0752 9329	.0758 6421	.0764 3744	.0770 1299	.0775 9087	.0781 7107	.0787 5361	.0793 3847	.0799 2567
.28	.0805 1521	.0811 0709	.0817 0131	.0822 9787	.0828 9679	.0834 9805	.0841 0166	.0847 0763	.0853 1596	.0859 2665
.29	.0865 3971	.0871 5513	.0877 7292	.0883 9308	.0890 1562	.0896 4054	.0902 6783	.0908 9751	.0915 2957	.0921 6403
.30	.0928 0088	.0934 4012	.0940 8176	.0947 2579	.0953 7224	.0960 2109	.0966 7234	.0973 2601	.0979 8210	.0986 4060
.31	.0993 0153	.0999 6488	.1006 3065	.1012 9886	.1019 6950	.1026 4257	.1033 1809	.1039 9605	.1046 7645	.1053 5930
.32	.1060 4461	.1067 3237	.1074 2259	.1081 1527	.1088 1041	.1095 0802	.1102 0810	.1109 1066	.1116 1569	.1123 2321
.33	.1130 3321	.1137 4569	.1144 6067	.1151 7814	.1158 9811	.1166 2057	.1173 4554	.1180 7303	.1188 0302	.1195 3552
.34	.1202 7054	.1210 0808	.1217 4815	.1224 9074	.1232 3587	.1239 8353	.1247 3373	.1254 8647	.1262 4175	.1269 9959
.35	.1277 5997	.1285 2292	.1292 8842	.1300 5649	.1308 2712	.1316 0032	.1323 7610	.1331 5446	.1339 3539	.1347 1891
.36	.1355 0503	.1362 9373	.1370 8503	.1378 7893	.1386 7544	.1394 7455	.1402 7628	.1410 8062	.1418 8758	.1426 9716
.37	.1435 0937	.1443 2421	.1451 4169	.1459 6180	.1467 8456	.1476 0997	.1484 3802	.1492 6873	.1501 0210	.1509 3813
.38	.1517 7683	.1526 1821	.1534 6325	.1543 0898	.1551 5838	.1560 1048	.1568 6527	.1577 2275	.1585 8293	.1594 4582
.39	.1603 1142	.1611 7973	.1620 5075	.1629 2450	.1638 0097	.1646 8017	.1655 6211	.1664 4679	.1673 3421	.1682 2437
.40	.1691 1729	.1700 1296	.1709 1140	.1718 1259	.1727 1657	.1736 2331	.1745 3283	.1754 4514	.1763 6023	.1772 7811
.41	.1781 9879	.1791 2228	.1800 4857	.1809 7767	.1819 0958	.1828 4432	.1837 8188	.1847 2227	.1856 6550	.1866 1156
.42	.1875 6047	.1885 1223	.1894 6684	.1904 2430	.1913 8464	.1923 4784	.1933 1391	.1942 8287	.1952 5470	.1962 2942
.43	.1972 0704	.1981 8755	.1991 7097	.2001 5730	.2011 4654	.2021 3869	.2031 3377	.2041 3178	.2051 3272	.2061 3661
.44	.2071 4343	.2081 5321	.2091 6594	.2101 8162	.2112 0028	.2122 2191	.2132 4651	.2142 7409	.2153 0466	.2163 3822
.45	.2173 7478	.2184 1434	.2194 5691	.2205 0249	.2215 5110	.2226 0272	.2236 5738	.2247 1508	.2257 7581	.2268 3960
.46	.2279 0643	.2289 7633	.2300 4929	.2311 2532	.2322 0442	.2332 8661	.2343 7189	.2354 6025	.2365 5172	.2376 4629
.47	.2387 4397	.2398 4477	.2409 4869	.2420 5574	.2431 6593	.2442 7925	.2453 9573	.2465 1535	.2476 3813	.2487 6409
.48	.2498 9320	.2510 2550	.2521 6099	.2532 9966	.2544 4152	.2555 8659	.2567 3487	.2578 8636	.2590 4107	.2601 9902
.49	.2613 6019	.2625 2462	.2636 9227	.2648 6319	.2660 3736	.2672 1481	.2683 9607	.2695 7952	.2707 6680	.2719 5737
.50	.2731 5125	.2743 4843	.2755 4892	.2767 5273	.2779 5987	.2791 6529	.2803 8415	.2816 0132	.2828 2183	.2840 4570
.51	.2852 7295	.2865 0356	.2877 3756	.2889 7495	.2902 1573	.2914 5992	.2927 0752	.2939 5853	.2952 1297	.2964 7084

The above theory applies to the quarter wave line. This is a rather difficult practical one; it is best to add another quarter wave to make a half-wave resonator, shorted at both ends, with the dielectric positioned exactly in the center. From conditions of symmetry we then employ the above equations, taking half of our measured quantities. Or, in terms of the actually measured four lengths which constitute an observation on a half-wave line, $(d-d')$, t, $\Delta\ell$ and λ, we have,

$$\left.\begin{aligned} \text{P.F.} &= \frac{\sinh \frac{\pi}{\lambda}(d-d')}{\sin \frac{2\pi}{\lambda}(\Delta\ell + t)} \cdot \frac{\sin 2X}{2X} \\ \frac{\pi t}{\lambda} \cdot \tan \pi \frac{\Delta\ell + t}{\lambda} &= X \tan X \\ \epsilon &= \left[\frac{X}{\frac{\pi t}{\lambda}}\right]^2 \end{aligned}\right\} \tag{15}$$

which are the expressions used in this work.

In practice the dielectric plug is pushed into the half-wave line and the line is tuned. The line center is then calculated and the plug reset to this. Retuning checks the correct location. Two trials are always sufficient if the plug was nearly centered originally.

There are several shortcomings affecting this theory. The Q of the unloaded line depends partly on metal power loss along the line. When the line is shortened by the dielectric plug, part of this loss disappears and part is transferred to the dielectric plug. Fortunately these losses are small since they are metal losses at a current node, but for long dielectric plugs or plugs of high dielectric constant the need for correction can arise. The necessary calculations have not yet been reduced to a simple form.

Again, the calculation of half-wave results by means of a quarter wave theory is safe only for a high Q situation. It is easy to show, experimentally, that the maximum line shortening results when the dielectric plug is exactly centered in the line but the calculated power factor is not a maximum here, as might be expected. In the meantime, experience shows that results can be duplicated from day to day and at other frequencies and that over a reasonable range of plug thickness no change in dielectric constant and power factor values, greater than the unavoidable errors of measurement, is obtained.

Description of Apparatus

The apparatus can be divided into three parts for purposes of description. The high frequency generator consists of a small "relay rack" assembly,

including 60-cycle power panel, rectifier panel, meter and control panel and centimeter wave oscillator panel with coaxial conductor output jack. All high-frequency connectors are coaxial conductor units with plug tips.

The measuring unit is shown in the two photographs; Fig. 1, assembled and Fig. 2, disassembled. Two combination input-output heads are shown in Fig. 2. These heads and tubing together with center conductor and plunger are of coin silver. While the highest possible conductivity metal is desirable, pure silver is mechanically too poor for spring fingers and bearing surfaces and the alloy must be used. The good sliding contact properties of silver are preserved but the conductivity is no better than that of copper. Both heads are drilled, for input and output connections, flush with the bottom of the cylindrical cavity terminating the tubing.

Head #1, shown attached in Fig. 1 and detached in lower right-hand corner of Fig. 2, has a silicon crystal, mounted and insulated in a small cylindrical holder which carries a tiny pickup loop, one side of which is grounded to the cylinder. The total length of pickup conductor including loop and crystal "whisker" is about one centimeter and no tuning is necessary. The loop pickup can be adjusted by moving the holder in or out. The d-c circuit is from an insulated pin on the holder through crystal to apparatus body.

The current input connection is through a coaxial plug which is tapped across a fraction of a tunable half-wave line. This fraction consists of a $\frac{1}{8}''$ coaxial conductor terminated in a tiny feed loop; the remainder of the line is an ordinary $\frac{1}{4}''$ coaxial with sliding plunger. The line is used, well off tune, as an input current amplitude control. The coupling with the cavity in head is adjusted by moving the feed loop in or out.

By inverting another half-wave coaxial with feed loop, so as to put the crystal where the feed jack was, it is possible to use an externally mounted crystal as in head #2. For this head the input current amplitude control is obtained by using, as a feeder, a short $\frac{1}{8}''$ coaxial tipped with a tiny loop and a coaxial jack, at opposite ends. This coaxial is mounted in a spring clamped bearing so as to permit a rotation of the plane of the loop. All coaxials, except the measuring unit itself, are 72-ohm ones.

There is no essential difference in operation between these two heads; they are interchangeable. However, head #1 is more convenient in manipulation, during the disassembly required to insert the dielectric sample. (This sample is always positioned in the piece of tubing connecting to the head.)

An ordinary model 301 microammeter, low resistance, served as indicating instrument. By replacing the crystal holder of head #2 with a loop tipped coaxial and plug, a conventional double-detection radio receiver with

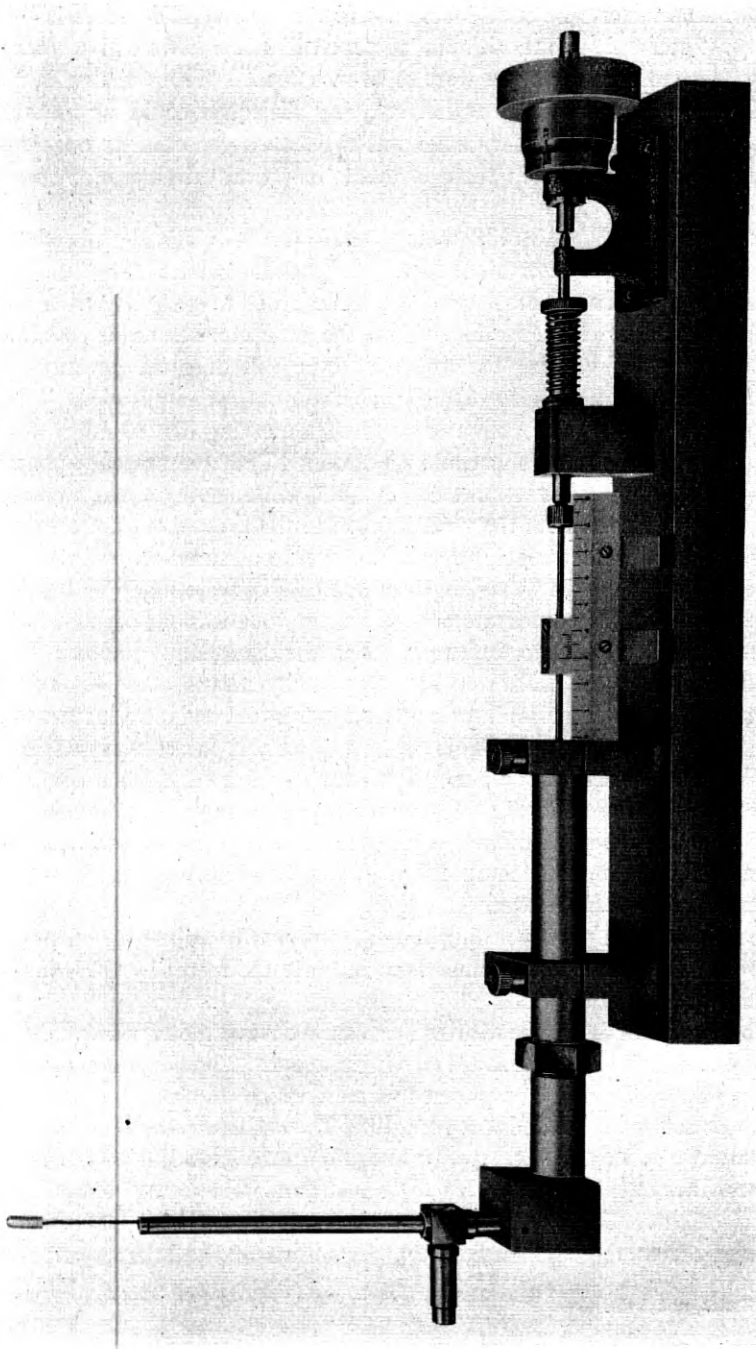

Fig. 1—Measuring unit, assembled

output meter could be used instead. The crystal type detector is by far the most convenient but with the power available wouldn't give workable outputs when bad dielectrics were to be measured. With the amplification available in the double detection set, any dielectric could be measured, while retaining the necessary attenuation between generator-resonator and resonator-receiver to keep these elements electrically independent of each other.

It is necessary to maintain an electrical isolation of this sort to get a high apparatus Q. The equivalent Q of all good dielectrics being high, the measuring apparatus Q must be of the same order to give favorable measuring conditions. And, further, unless the generator-resonator coupling is weak, the act of varying the resonator tune will drag the generator frequency around and will also vary the generator output amplitude.

The crystal plus microammeter required something like 80 millivolts for full scale deflection and this could be obtained with the present apparatus with couplings giving a resonator Q of 1500, while having enough power in reserve to measure any of the good dielectrics. However, most of the dielectrics with power factor greater than .01 were measured with the d.d. receiver. All the 10 cm wave-length measurements were made with this receiver. For the latter measurements a shorter tube was substituted for the tubes shown screwed into the two heads in the disassembly photo.

The crystals were calibrated at 60 cycles by means of a 70-ohm $\sqrt{2}$ attenuation pad.[2] With full scale deflection this pad was introduced and the new scale deflection read. This $\sqrt{2}$ ratio was, as far as was possible to check, maintained in the kilo megacycle range. For calibration the crystal was tapped across 4 ohms in the attenuator pad output. A 15 mf electrolytic condenser was permanently connected across the meter terminals and, by means of a pair of switches, calibration could be checked in a few seconds, during a measurement run.

The calibration process, using the d.d. set, was to adjust the output to a convenient meter deflection and then calibrate the meter by throwing in 3 db in the IF attenuator.

The resonator itself constitutes an accurate wave meter when corrected for the change in diameter at the moving plunger. The method of operation was then as follows. The plunger vernier, which allowed reading to 0.01 cm., was set at the desired wave-length. The oscillator was then turned on and after it had attained temperature equilibrium, was adjusted if necessary to resonance at this value. This adjustment was infrequently necessary and always slight. The apparatus Q was then determined by traversing the plunger across the resonance setting by means of the micrometer. This

[2] Exact, not 3 db.

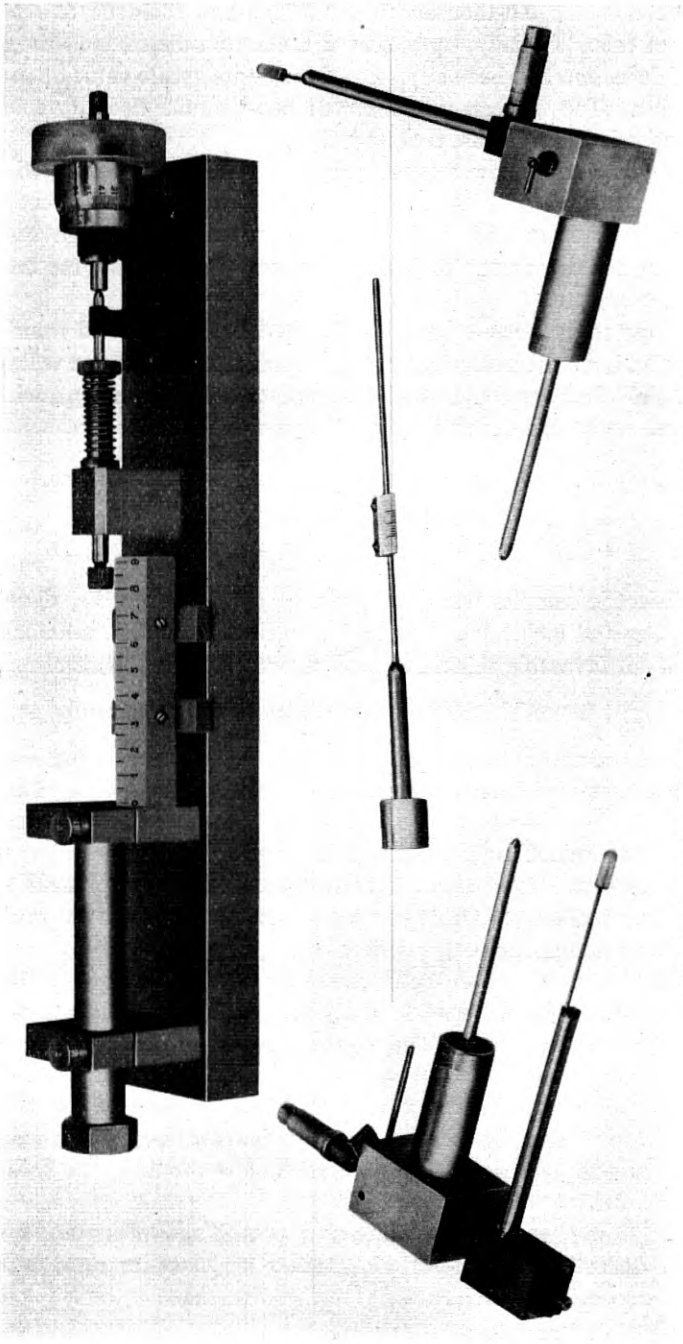

Fig. 2—Measuring unit, disassembled. Two different heads shown

"mike" read to the ten-thousandth of an inch and could be estimated to one-fifth of this. Initially, by means of the amplitude control, the microammeter deflection had been adjusted to the desired scale value at the resonance point. The traverse was observed between the two $\sqrt{2}$ microammeter deflections and was repeated in the opposite direction. When successive round trips showed consistency the value of d' was noted. The dielectric sample, after thickness measurement, was then introduced, centered by cut and try and the Q traverses repeated. This gave d and, after noting $\Delta\ell$, the change in plunger setting for resonance, the measurement was complete.

During the measurement the generator had to be protected from drafts and, usually, it was necessary to traverse rapidly, the power line voltage not being stable. Settings could usually be reproduced to 1 per cent, with adequate care. A sample observation on a good dielectric is the following:

July 28, 1941 Polystyrene plate, all dimensions in cms.

$t = 1.28$	$d' = .0084$	$\lambda = 22.42$
$\Delta\ell = 1.79$	$d = .010$	P.F. $= .00028$, $\epsilon = 2.49$

The dielectric samples were machined on a precision lathe, dimensions being held to .001 inch. The nominal dimensions were O.D. .640 inch, I.D. .174 inch. A favorable thickness, from the standpoint of ease of measurement, is $\left|\frac{\lambda}{10\epsilon}\right|$, in cm's. Cleanliness in handling was carefully observed. After a lapse of several days the interior bearing surfaces of the resonance cavity would have to be cleaned with fine French crocus cloth. The plunger bearing surfaces also had to be smoothed up, fine scratches being polished off. Dirt was immediately noticeable when the plunger contacted it, and when microscopic bits of silver were rolled up under the plunger springs cleaning was necessary. Otherwise no particular treatment or smoothing up of the contacting surfaces was required.

A table of dielectric power factors and constants is a very desirable piece of information. Unfortunately, experience tends to the conclusion that such a table does not exist. The organic plastics in particular, are rather variable from sample to sample and a table of values is merely a table for particular specimens. Where a great number of samples are available "best", "worst" and "most common" values can be established. The accompanying list of observed values must be interpreted in the light of the above statements.

As a large number of measurements of certain special materials had to be made, dielectrics in general were rather neglected and the tabulated values are more or less incidental. It was noted that for the low loss, sub-

TABLE 1

Material	ε		P.F.	
	22.5 cms.	10 cms.	22.5 cms	10 cms.
Ceramic				
BTL F3 Mg. Silicate type	5.83		.00023	
"Dielectene"		3.39		.0038
Glass, Corning				
G1, lead	4.30		.0049	
G8, lime, annealed	6.38		.0102	
G12, lead	6.08		.0035	
199-1	8.70		.0019	
702EJ, Pyrex	6.35		.0067	
702P	4.70		.0053	
704EO	4.42		.0033	
705BA	3.80		.00118	
707DG	4.69	4.8	.0037	.0036
Glyptal	3.38	3.36	.030	.036
Lucite	2.58	2.56	.0090	.0087
Mycalex				
Red	5.91		.0030	
White	5.74		.0033	
Phenolics				
Cast specimen		4.63		.139
Bakelite sheet $\frac{1}{32}$"		3.57		.080
Polyethylene				
Worst			.00229	
Most common	2.26		.00060	
Best			.00031	
Polystyrene				
Worst			.00090	
Most common	2.45		.00070	
Best			.00028	
Polyvinylcarbazole	2.87		.0040	
Rubber				
Hard, brown		2.77		.0041
Hard, black		2.69		.0059
Soft, black	3.15		.0058	
Resin	2.32		.0018	
Styralloy				
No. 10	2.49		.0036	
Desig. Unknown	2.49	2.50	.0019	.00105
No. 22	2.40		.0047	
Styramic				
E1689		2.55		.00087
Tenite II		2.95		.031
Vinylite V	2.78	2.61	.0076	.0068
Wax				
Paraffin	2.17		.00019	
Boler	2.17		.00019	
Superla	2.26	2.26	.00019	.00015

stituted paraffin-type, carbon chain dielectrics no difference, greater than experimental error, exists between the 22.5 and 10 cm. measurements.

ACKNOWLEDGEMENT

The measurements by means of the double-detection set were made by my co-worker, Mr. W. E. Eckner, whose valuable assistance I am glad to acknowledge. To Mr. C. F. Mattke, also of the Bell Laboratories, I am indebted for assistance in getting my crude original apparatus into its final finished form.

APPENDIX

The typical ultra high-frequency transmission line can be represented as in Fig. 3

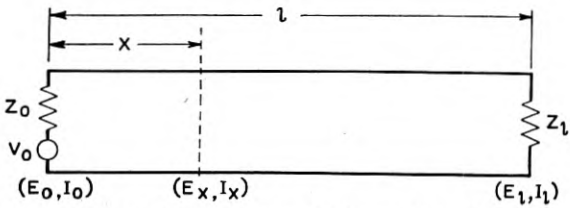

Fig. 3—Equivalent circuit of transmission line

and the equations describing it are

$$\left.\begin{aligned}
E_x &= V_0 \frac{Z_\ell \cosh \sqrt{DZ}\,(\ell - x) + \sqrt{\frac{Z}{D}} \sinh \sqrt{DZ}\,(\ell - x)}{(Z_0 + Z_\ell)\cosh \sqrt{DZ}\,\ell + \left(Z_0 Z_\ell \sqrt{\frac{D}{Z}} + \sqrt{\frac{Z}{D}}\right)\sinh \sqrt{DZ}\,\ell} \\
I_x &= V_0 \frac{\cosh \sqrt{DZ}\,(\ell - x) + Z_\ell \sqrt{\frac{D}{Z}} \sinh \sqrt{DZ}\,(\ell - x)}{(Z_0 + Z_\ell)\cosh \sqrt{DZ}\,\ell + \left(Z_0 Z_\ell \sqrt{\frac{D}{Z}} + \sqrt{\frac{Z}{D}}\right)\sinh \sqrt{DZ}\,\ell} \\
Z_x &= \frac{E_x}{I_x} = \sqrt{\frac{Z}{D}} \cdot \frac{Z_\ell + \sqrt{\frac{Z}{D}} \tanh \sqrt{DZ}\,(\ell - x)}{\sqrt{\frac{Z}{D}} + Z_\ell \tanh \sqrt{DZ}\,(\ell - x)} \\
E_0 &= V_0 - Z_0 I_0\,, \qquad E_\ell = Z_\ell I_\ell
\end{aligned}\right\} \quad (1)$$

The line constants are $Z = R + i\omega L$, the series impedance per unit length, and $D = G + i\omega C$, the shunt admittance per unit length. From these we have: surge impedance $= \sqrt{\frac{Z}{D}} = S_0$, propagation constant $= \sqrt{DZ}$.

For all lines usable as transmission devices the following approximations hold:

$$\sqrt{DZ} = \alpha + i\beta \quad \alpha = \frac{R}{2}\sqrt{\frac{C}{L}} + \frac{G}{2}\sqrt{\frac{L}{C}} \quad \beta = \omega\sqrt{LC} = \frac{\omega}{v} = \frac{2\pi}{\lambda}$$
$$\sqrt{\frac{Z}{D}} = \sqrt{\frac{L}{C}}, \quad C = \frac{1}{S_0 V}, \quad L = \frac{S_0}{V}, \quad v = 3 \times 10^{10} \text{ cm/sec. for air.} \tag{2}$$

For the case of near-end input and output the second of equations (1) can be rewritten as

$$I_0 = \frac{V_0}{\sqrt{\frac{Z}{D}}} \cdot \frac{\cosh \sqrt{DZ}\,\ell + \frac{Z_\ell}{\sqrt{\frac{Z}{D}}} \sinh \sqrt{DZ}\,\ell}{\frac{Z_\ell}{\sqrt{\frac{Z}{D}}} \cosh \sqrt{DZ}\,\ell + \sinh \sqrt{DZ}\,\ell + \frac{Z_0}{\sqrt{\frac{Z}{D}}}\left(\cosh \sqrt{DZ}\,\ell + \frac{Z_\ell}{\sqrt{\frac{Z}{D}}} \sinh \sqrt{DZ}\,\ell\right)} \tag{3}$$

If now we assume a quarter-wave line, the condition of resonance implies that $R_\ell \gg \left|\sqrt{\frac{Z}{D}}\right|$ and $R_0 \ll \left|\sqrt{\frac{Z}{D}}\right|$. The condition of reasonable shortening of the line (or lengthening) by the terminal reactance implies that $X_\ell > \left|\sqrt{\frac{Z}{D}}\right|$, $X_0 < \left|\sqrt{\frac{Z}{D}}\right|$. Hence we shall have $|Z_\ell| \gg \left|\sqrt{\frac{Z}{D}}\right|$, $|Z_0| \ll \left|\sqrt{\frac{Z}{D}}\right|$.

If we put $\frac{\sqrt{\frac{Z}{D}}}{Z_\ell} = \tanh \theta$, we get

$$I_0 = \frac{V_0}{\sqrt{\frac{Z}{D}}} \cdot \frac{\tanh\left(\sqrt{DZ}\,\ell + \theta\right)}{1 + \frac{Z_0}{\sqrt{\frac{Z}{D}}} \tanh\left(\sqrt{DZ}\,\ell + \theta\right)} \tag{4}$$

We now make the assumption that "Z_0" is a pure resistance (which is no limitation on the measurement to be discussed) and put $\theta = a_\ell + ib_\ell$.

Then,

$$|I_0| = \frac{V_0}{\sqrt{\frac{L}{C}}} \cdot \sqrt{\frac{\tanh^2(\alpha\ell + a_\ell) + \tan^2\left(\frac{2\pi\ell}{\lambda} + b_\ell\right)}{\left[1 + \frac{Z_0}{\sqrt{\frac{L}{C}}} \tanh(\alpha\ell + a_\ell)\right]^2 + \left[\frac{Z_0}{\sqrt{\frac{L}{C}}} + \tanh(\alpha\ell + a_\ell)\right]^2 \tan^2\left(\frac{2\pi\ell}{\lambda} + b_\ell\right)}} \tag{5}$$

This expression cycles, as "ℓ" is varied, and has its maximum or "tuned" value of

$$\frac{V_0 \Big/ \sqrt{\frac{L}{C}}}{\frac{Z_0}{\sqrt{\frac{L}{C}}} + \tanh(\alpha\ell + a_\ell)} \quad \text{for} \quad \tan\left(\frac{2\pi\ell}{\lambda} + b_\ell\right) = \infty$$

$$\text{or} \quad \frac{2\pi\ell}{\lambda} + b_\ell = \frac{(2n+1)\pi}{2} \qquad n = 0, 1, 2, \cdots$$

The resonant length is thus $\ell = \frac{\lambda}{4}\left(1 - \frac{2b_\ell}{\pi}\right)$, for $n = o$. Note that successive resonances differ by a line length of $\frac{\lambda}{2}$; the reactive termination has merely shortened, by the amount of $\Delta\ell = \frac{\lambda b_\ell}{2\pi}$, the first resonant element preceding it. When, therefore, we measure the "Q" of this line segment by line-length tuning we use $\ell = \frac{\lambda}{4}$ in the Q process definition.

The Q process now follows. Putting $\ell = \ell_r \pm \delta\ell$ where ℓ_r is the actual observed resonance length, we have

$$\frac{2\pi\ell}{\lambda} + b_\ell = \frac{2\pi\ell_r}{\lambda} + b_\ell \pm \frac{2\pi\delta\ell}{\lambda} = \frac{\pi}{2} \pm \frac{2\pi\delta\ell}{\lambda}$$

Then $\tan\left(\frac{2\pi\ell}{\lambda} + b_\ell\right) = \tan\left(\frac{\pi}{2} \pm \frac{2\pi\delta\ell}{\lambda}\right) = \frac{1}{\mp \tan\frac{2\pi\delta\ell}{\lambda}}$ and

$$|I_0| = \frac{V_0}{\sqrt{\frac{L}{C}}} \sqrt{\frac{1 + \tanh^2(\alpha\ell + a_\ell)\cdot\tan^2\frac{2\pi\delta\ell}{\lambda}}{\left[\frac{Z_0}{\sqrt{\frac{L}{C}}} + \tanh(\alpha\ell + a_\ell)\right]^2 + \left[1 + \frac{Z_0}{\sqrt{\frac{L}{C}}}\tanh(\alpha\ell + a_\ell)\right]^2 \tan^2\frac{2\pi\delta\ell}{\lambda}}}$$

Forming the current values $|I_{01}| = |I_{02}| = \left|\frac{I_0(\text{resonant})}{K}\right|$, dividing to eliminate $|I_0|$ and discarding squares and products of small quantities in comparison with unity leaves,

$$\sqrt{K^2 - 1} = \frac{\tan\frac{2\pi\delta\ell}{\lambda}}{\frac{Z_0}{\sqrt{\frac{L}{C}}} + \tanh(\alpha\ell + a_\ell)}$$

$$\text{or} \quad \sqrt{K^2 - 1}\,\frac{\lambda}{8\delta\ell} = \frac{\frac{\lambda}{8\delta\ell}\cdot\tan\frac{2\pi\delta\ell}{\lambda}}{\frac{Z_0}{\sqrt{\frac{L}{C}}} + \tanh(\alpha\ell + a_\ell)} \tag{6}$$

which becomes our "Q" when $K = \sqrt{2}$.

In most practical situations the "tan" and "tanh" are equal to their angles. For this condition $Q = \frac{\frac{\pi}{4}}{\frac{Z_0}{\sqrt{\frac{L}{C}}} + \alpha\ell + a_\ell}$. If we now put, by definition, $Q_\ell = \frac{\pi}{4\alpha\ell}$ (the Q of the line itself), $Q_0 = \frac{\pi}{4}\cdot\frac{\sqrt{\frac{L}{C}}}{Z_0}$, $Q_\ell = \frac{\pi}{4a_\ell}$, we have

$$\frac{1}{Q} = \frac{1}{Q_\ell} + \frac{1}{Q_0} + \frac{1}{Q_\ell} \tag{7}$$

the law of Q composition relating the resultant Q to line and terminal Q's.

Abstracts of Technical Articles by Bell System Authors

A Sampling Inspection Plan for Continuous Production.[1] H. F. DODGE. This paper presents a plan of sampling inspection for a product consisting of individual units (parts, sub-assemblies, finished articles, etc.) manufactured in quantity by an essentially continuous process.

The plan, applicable only to characteristics subject to non-destructive inspection on a Go-NoGo basis, is intended primarily for use in process inspection of parts or final inspection of finished articles within a manufacturing plant, where it is desired to have assurance that the percentage of defective units in accepted product will be held down to some prescribed low figure. It differs from others which have been published in that it presumes a *continuous flow* of *consecutive articles* or *consecutive lots* of articles offered to the inspector for acceptance in the order of their production. It is accordingly of particular interest for products manufactured by conveyor or other straight line continuous processes.

In operation, the plan provides a corrective inspection, serving as a partial screen for defective units. Normally, a chosen percentage or fraction f of the units are inspected, but when a defective unit is disclosed by the inspection it is required that an additional number of units be inspected, the additional number depending on how many more defective units are found. The result of such inspections is to remove some of the defective units, and the poorer the quality submitted to the inspector, as measured in terms of per cent defective, the greater will be the corrective or screening effect. The object of the plan is the same as that incorporated in some of the sampling tables already published, namely, to establish a limiting value of "average outgoing quality" expressed in per cent defective which will not be exceeded no matter what quality is submitted to the inspector. This limiting value of per cent defective is termed the "average outgoing quality limit (AOQL)."

The theoretical solution treats the case of inspecting a continuous flow of individual units and is based on the distribution of *random-order* spacing of defective units in product whose quality is statistically controlled. Part III of the paper extends the application of the method to a continuous flow of individual lots or sub-lots of articles.

Stability in High-Frequency Oscillators.[2] R. A. HEISING. This paper discusses frequency stability with change in plate voltage of high-frequency

[1] *The Annals of Mathematical Statistics*, September 1943.
[2] *Proc. I.R.E.*, November 1943.

oscillators of around 100 megacycles and shows both theoretically and experimentally that the highest stability found by many is only the result of fortuitous circuit adjustment that may readily lead to the desired result in this frequency range. It is shown that the factor next in importance in producing frequency stability is a low ratio of inductance to capacitance in the frequency-determining circuit. It is also shown that a high Q contributes little directly to stability. A high Q is necessary with low L/C ratios to get oscillations but an improvement in Q alone may give poorer stability. To get the fullest measure of stability with low L/C and high Q calls for slight adjustments in the circuit and possibly the provision of loose coupling to the frequency-determining circuit.

Modern Spectrochemical Analysis.[3] EDWIN K. JAYCOX. The spectrograph, originally developed by the physicist, has become a most useful tool in the hands of the analytical chemist. Today few large analytical laboratories are without one. The instrument, with its attendant accessories, provides a rapid method for analyzing metals, alloys, minerals, ores, liquids, and gases, particularly for their metallic constituents and in some cases for their anions. Both emission and absorption spectra are important to the analyst. Important applications of the spectrograph to the analytical problems of research and industrial organizations are discussed.

The spectrograph did not come into general use as an analytical tool until the early 1920's, although Kirchhof and Bunsen saw the practicability of the method in 1860, when they published their paper entitled, "Chemical Analysis by Means of Spectral Observations." During the intervening years only a few enthusiasts like Lockyer, Roberts, Hartley, Leonard, Pollack, and de Gramont, kept the art alive. In spite of their persistent efforts to influence chemists to use spectrographic methods, they were quite generally ridiculed and the value of the method was recognized by only a few workers.

In 1922, Meggers, Kiess, and Stimson published their paper "Practical Spectrographic Analysis" and modern spectrochemical analysis was born. Under the stimulus of this paper and the backing of a high caliber scientific organization like the Bureau of Standards, the use of the spectrograph as an analytical tool increased rapidly. This is evidenced from the *Index to the Literature on Spectrochemical Analysis* by Meggers and Scribner. In 1920, for example, only five papers were published concerning spectrochemical analysis, four of which were by de Gramont; whereas in 1930, 33 papers were published and in 1939, 170 papers, indicating an increasing interest in and use of spectrochemical analysis in industrial and research organizations.

[3] *Jour. Applied Physics,* December 1943.

Determination of Small Amounts of Arsenic, Antimony, and Tin in Lead and Lead Alloys.[4] C. L. LUKE. A new method for the determination of small amounts of arsenic, antimony, and tin in lead and lead alloys consists of separation of the three metals from the lead by a double co-precipitation with manganese dioxide, reduction of arsenic and antimony to the trivalent state, separation of the arsenic by distillation as chloride, titration of the arsenic and antimony separately by the method of Gyory, and reduction of tin with lead and titration with standard iodine solution.

Determination of Total Sulfur in Rubber.[5] C. L. LUKE. A new volumetric method has been developed for the determination of sulfate sulfur. The sulfate is reduced to sulfide by treatment with hydriodic acid and the hydrogen sulfide is distilled off and titrated iodometrically. The new method has been applied to the determination of total sulfur in natural and synthetic rubber.

Machine Screws. Fastening Strengths in Various Materials.[6] A. C. MILLARD. Although standard machine screws in the numbered sizes have been widely used as fastenings for many years, very little has been published concerning their strength of fastening in various metals and non-metals. Numerous articles have appeared regarding the strength of bolts and machine screws for $\frac{1}{4}$ in. and larger sizes, but very little, if any, published information is available on the strength of machine-screw fastenings in the numbered sizes.

The need for machine-screw fastening-strength information has increased recently due to the use of more compact designs and the shortage of materials. The use of substitute materials has accentuated the lack of machine-fastening-strength information in making fastenings in such materials, as well as in the more commonly used materials. Frequently, it is desirable to know the load-carrying capacity of screw fastenings of various diameters, as well as the length of thread engagement in the weaker materials needed to develop either the full strength of the screw, or the strength of fastening required of the assembly. The purpose of this paper is to make available to designers the results of fastening-strength tests of machine-screw fastenings in a number of materials, which were carried out by the author at the Bell Telephone Laboratories, Inc. The work is by no means complete but is hoped that the data offered will prove to be of some use in its present form.

[4] *Indus. & Engg. Chemistry*, October 1943.
[5] *Indus. & Engg. Chemistry*, September 1943.
[6] *Mech. Engg.* October 1943.

Contributors to this Issue

W. R. BENNETT, B.S., Oregon State College, 1925; A.M., Columbia University, 1928. Bell Telephone Laboratories, 1925–. Mr. Bennett has been engaged in the study of the electrical transmission problems of communication.

CARL R. ENGLUND, B.S. in Chemical Engineering, University of South Dakota, 1909; University of Chicago, 1910–12; Professor of Physics and Geology, Western Maryland College, 1912–13; Laboratory Assistant, University of Michigan, 1913–14. Western Electric Company, 1914–25; Bell Telephone Laboratories, 1925–. As radio research engineer Mr. Englund is engaged largely in experimental work in radio communication.

D. K. GANNETT, B.S. in engineering, University of Minnesota, 1916; E.E. University of Minnesota, 1917. American Telephone and Telegraph Company, Engineering Department, 1917–1919; Department of Development and Research, 1919–1934; Bell Telephone Laboratories, Inc. 1934–. Prior to October 1942 Mr. Gannett, as Toll Transmission Engineer, was concerned with the transmission requirements of toll systems including program circuits. Since then, as Circuit Research Engineer, he has directed a group engaged in research and development on war projects.

IDEN KERNEY, B.S. Harvard University, 1923. American Telephone and Telegraph Company, 1923–1934; Bell Telephone Laboratories, Inc. 1934–. Before the war Mr. Kerney was in charge of the laboratory in which experimental work on program transmission was conducted. Since early in 1942 he has been engaged full time on war projects.

R. A. SYKES, Massachusetts Institute of Technology, B.S. 1929; M.S. 1930. Columbia University, 1931–1933. Bell Telephone Laboratories, Research Department, 1930–. Mr. Sykes has been engaged in the applications of quartz crystals to broad-band carrier systems as filter and oscillator elements. Other work has included the application of coaxial lines as elements of filter networks and more recently the design and development of quartz crystals for radio frequency oscillators.

G. W. WILLARD, B.A., University of Minnesota, 1924; M.A., 1928; Instructor in Physics, University of Kansas, 1927–28; Student and Assistant, University of Chicago, 1928–30. Bell Telephone Laboratories, 1930–. Mr. Willard's work has had to do with special problems in piezo-electric crystals.