

Introduction

HISTORICAL SURVEY

The early history of the subject is covered in detail in Cady's classic work *Piezoelectricity* (Cady, 1964), and in Heising (1946). The direct piezoelectric effect, the production of electricity by the application of pressure, was discovered by the brothers J. and P. Curie in 1880. The converse effect was predicted by Lippmann in 1881 and confirmed by the Curies in the same year. The first major application of piezoelectricity was in Langevin's work on submarine detection using quartz transducers to generate and detect underwater acoustic waves, started in 1917. Nicolson in 1919 described several devices, such as loudspeakers, microphones and sound pick-ups based on the piezoelectric properties of Rochelle salt, and this was followed by Cady's publication in 1921 of papers describing the first quartz crystal controlled oscillator. From this time considerable research effort was devoted to the field of quartz crystals.

In 1926, the station WEAf in New York became the first broadcast transmitter to use quartz crystal control. At first the further application of crystal control in radio communications was hampered by the temperature dependence of the crystal frequency, which made necessary the use of cumbersome temperature control equipment, especially in applications such as aircraft communications. This obstacle was effectively removed by the introduction in 1934 of crystal cuts with very low temperature coefficients of frequency (Lack *et al.*, 1934). The early work on frequency control was paralleled by work on frequency selective devices, and the quartz crystal filter had become an essential component of telephone line transmission systems by 1939.

With the approach of World War II, it was recognized by the military authorities that crystal control was essential if the full benefit of radio communications was to be obtained in battle. The consequent demand for quartz crystals resulted in a tremendous expansion of the then embryo crystal industry. In the United Kingdom, production increased from 10 000 units in 1938 to 1.5 million units in 1945 (Vigoreux and Booth, 1950, p. 1), and in the USA the Western Electric company's crystal production increased 750-fold during the war period (Heising, 1946, p. 8). This expansion was only possible thanks to the concerted efforts of many people in solving the many design and manufacturing problems that were encountered, and the subsequent

dissemination of this knowledge in such works as those of Heising and of Vigoreux and Booth, helped to lay the foundation for the subsequent growth of the industry.

The post-war period has seen sustained development in the traditional application areas of frequency control and selection in radio and line communication systems. Shortages of natural quartz during the war years inspired both a search for alternative piezoelectric materials and also research into techniques for producing cultured quartz crystals. The success of the latter program is evident insofar as cultured quartz, commercially available from several sources, has now practically replaced natural quartz as the industry's raw material. Improvements in design methods and manufacturing techniques have produced successive advances in crystal unit performance, particularly in regard to miniaturization, increased frequency stability, and increased frequency range. Crystal filters, previously mainly restricted to line transmission systems, have been increasingly used since the late 1950s in other applications, particularly as channel filters in mobile radio systems. This trend was accelerated by the introduction of the monolithic crystal filter in the middle 1960s.

Particularly in the case of radio communications, the historical demand for quartz crystals has for the most part been characterized by requirements for many different frequencies, with relatively small quantities on any particular frequency. The typical end use was that of a multi-channel radio set requiring two crystals per channel, one for transmit, one for receive. Generally, different users operated on different channel frequencies, and therefore required 'channel' crystals unique to themselves. Consequently the crystal industry developed with a manufacturing philosophy oriented to small batch production in an environment where customer service, in terms of rapid response to a customer's specific demands, was paramount.

In recent years this orientation has proved increasingly inappropriate. The impact of frequency synthesis techniques has reduced the overall requirement for channel crystals, and substituted a growing demand for stable reference frequency sources that are in principle better suited for standardization. Coupled with the demands for greater frequency stability resulting from the use of narrow channel spacings at higher and higher channel frequencies, the modern trend is towards the use of packaged temperature-compensated crystal oscillators in synthesized systems. Just as 'de facto' industry standards have developed for mobile radio channel filters on frequencies of 10.7 and 21.4 MHz, so it is to be expected that a similar standardization of packaged oscillator specifications will come about. Emphasis in manufacturing will then have to be placed on achieving high quality and reliability at low cost, with the premiums presently available for rapid delivery of custom products such as channel crystals an anachronism.

These considerations are even more appropriate to the many new applications for quartz crystals that have developed in recent years. Crystals for

colour television, for clocks and watches, for video tape recorders and for all kinds of timing applications such as microprocessor clocks, are all required in high volumes and on fixed frequencies. So too are the simple packaged oscillators generically known as 'clock' oscillators that are tending to take the place of discrete crystals in the timing applications just mentioned. These products are essentially mass-produced commodity items, representing a major, fundamental, change from the traditional crystal products business. A further feature of the transformation of the crystal market brought about by these new commodity products is that in pursuing the objective of miniaturization, specifically in the case of crystals for electronic watches, photolithographic manufacturing techniques have been introduced that have in turn spawned a whole range of new miniature crystal units with applications far beyond that originally envisaged for watches.

In summary, in the century or so since the discovery of piezoelectricity, the history of quartz crystal devices can be divided roughly into five periods. In the first period, from 1880 to about 1920, piezoelectricity remained a scientific curiosity. The second period from 1920 to 1939 can be described as the classical period, when the foundations were being laid for the forced growth of the third period, World War II. The 20 years immediately following the war, up to about 1965, constituted a fourth period of steady growth and orderly development, and the fifth period, from 1965 to the present has seen a succession of fundamental changes in both markets and techniques which can fairly justify the epithet 'revolutionary'.

PIEZOELECTRICITY

Piezoelectricity is literally 'pressure electricity', the prefix *piezo* being derived from the Greek 'to press'. The direct piezoelectric effect discovered in 1880 by the brothers Curie refers to the electric polarization of certain materials brought about by the application of a mechanical stress. The converse effect, predicted by Lippman and confirmed by the Curies in 1881, refers to the deformation produced in the same materials by the application of an electric field. Details of these early researches are given by Cady (1964).

All dielectric materials exhibit the phenomenon of *electrostriction*, that is they are deformed when placed in an electric field. Piezoelectric materials are generally distinguished from purely electrostrictive ones by the relatively much larger piezoelectric deformation and by the fact that the piezoelectric deformation is reversible. Thus if, for example, a rod of quartz is cut in such a way that an applied field causes an elongation of the rod, reversing the direction of the field will cause the rod to contract, whereas in a non-piezoelectric material, whatever deformation is caused will be independent of the

direction of the field. Quantitatively, a piezoelectric deformation is essentially linearly dependent on the applied field, whereas an electrostrictive deformation depends on the square of the field. With the exception of ferroelectric materials and some piezoelectric ceramics, the electrostrictive effect is generally orders of magnitude smaller than the converse piezoelectric effect where the latter exists. This is certainly true of quartz.

The reversible nature of the piezoelectric effect implies that piezoelectric materials must be anisotropic, that is that their physical properties must depend on direction within the material. More precisely, such materials cannot have a centre of symmetry, for with a centre of symmetry, the reversal of an applied field could have no significance relative to the material's internal structure. An explanation of the origin of the piezoelectric effect in terms of an assumed molecular structure for quartz was first given by Lord Kelvin soon after the Curies' initial discovery, and although his proposed structure has since been proved incorrect by X-ray crystallography, his explanation can still be accepted as a qualitative and heuristic guide to understanding the phenomenon. Kelvin's explanation (Appendix 1) hinges on the assumption of distributions of charges on the molecular level which are such that certain deformations of the structure bring about a separation of the centres of gravity of the positive and negative charges and therefore the production of an electric dipole moment or electrical polarization.

Physically, the polarization in a piezoelectric material can be regarded as having two components, one directly related to the mechanical deformation of the material by some such mechanism as discussed above, and the second related to the electric field in the material by the normal mechanisms of dielectric polarization. These two components can in principle be isolated in suitable experiments. For example, if a sample of material was clamped in such a way that no mechanical deformation was possible, then the polarization of the sample would be entirely due to the applied electric field. Similarly, if the surface of the sample was coated with a conducting layer maintained at a constant electric potential so that the internal electric field was constrained to be zero, then the polarization of the sample would be entirely due to its deformation. In the general case, the polarization would be the sum of these two components, and be written in the form

$$P = kE + eS$$

where P , E and S are the polarization, electric field and mechanical strain, respectively, and k , e are, respectively, the susceptibility and the piezoelectric stress constant.

The deformation or strain in a piezoelectric material can likewise be thought of as having two components, the piezoelectric strain due to the electric field and the additional strain due to any mechanical stresses in the material. Again, these two components can in principle be separated in ideal experiments. Thus the piezoelectric strain would be that observed in the

material in the presence of an electric field and complete absence of any mechanical forces, whether body forces such as that due to gravity, or surface tractions. The purely elastic strain would be that due to applied stresses in a sample of material where the field was constrained to be zero. The general expression for the strain would then be written

$$S = sT + dE$$

where T is the stress in the material and s, d are, respectively, the elastic compliance and the piezoelectric strain constant.

It is tempting, but potentially confusing, to regard the piezoelectric strain as being the result of a piezoelectric stress due to the electric field. By definition, a stress is a force per unit area of a surface exerted by the material on one side of the surface and acting on the material on the other side. The piezoelectric strain is, however, caused by the electric field acting at long range and not by transmission of a stress through the body of the material. The distinction between the states of stress in an elastically deformed body as compared to a body similarly deformed by an electric field through the agency of the piezoelectric effect can be illustrated by considering the case of a long thin rod. Suppose first that such a rod is stretched by the application of equal and opposite longitudinal forces at its ends. The rod as a whole is in equilibrium and in a state of tensile stress, but if the rod is cut in two the equilibrium is destroyed. The two halves of the rod will revert to a state of zero strain and at the same time be accelerated by the unbalanced external forces acting on them. The situation is quite different in the case of the same rod stretched by the same amount by an electric field through the piezoelectric effect. Cutting the rod in this case now produces no effect whatsoever. The two halves remain in the same state of strain and remain stationary, thus indicating that there were no forces acting across the cross-section of the rod where the cut was made, that is, there is zero stress in the material.

Because of the tensor nature of the stress T and strain S and the necessarily anisotropic character of the material, it is not possible to interpret the material constants k, e, s and d as scalar quantities. P and E are vector quantities specified by their three components with respect to some coordinate system, so that the susceptibility k must be treated as a second rank tensor quantity. The stresses and strains T and S are themselves symmetric second rank tensors and so the piezoelectric constants e and d relating the vector quantities E and P to T and S must be treated as third rank tensors, and the elastic constants c and s are correspondingly fourth rank tensors. A brief discussion of tensor formalism and the formulation of the linear theory of piezoelectricity is given in the appendices.

ELASTIC VIBRATIONS IN SOLIDS

If a sample of elastic material is deformed by the application of external forces and those forces are suddenly removed, then vibrations will be set up in the material, gradually dying away as the energy stored in the material by the work done in the initial deformation is dissipated by the various internal and external loss mechanisms present. This phenomenon is familiar in such well known instances as bells, gongs, tuning forks and xylophones. In musical instruments, the pleasant effect of the resulting sound is due to the fact that the structure or geometry of the device results in the generation of a relatively small number of harmonically related frequencies. Historically, the design of such instruments as bells would have been arrived at empirically, and in general the problem of calculating the frequencies of a vibrating body in terms of its dimensions and material properties is very difficult. Nevertheless, the solution of the problem in the case of quartz crystal resonators is a necessary part of their design.

There are two situations to be considered. The first is the case of the free vibrations of a body set in motion by arbitrary forces and then released; the second the case of forced vibrations, where an external periodic driving force is present throughout the motion. In the former case, the vibrations vary from instance to instance depending upon the initial conditions; this is illustrated by the variations in tone that result from striking a bell in different positions on its surface, and can be explained physically in terms of the concept of normal modes of vibration. In a normal mode, all points of the body share in a (damped) simple harmonic motion and have the same phase and frequency. Thus when any one point goes through its equilibrium position, so too do all the other points, and when the same point reaches its extreme position, so too do the others. Analysis shows that there are generally an infinite number of such modes, and that an arbitrary free motion of the body can be regarded as a superposition of normal modes with different amplitudes and phases determined by the initial conditions. Practically speaking, it is impossible to excite a single normal mode, since this would require the initial deformation precisely to mirror that of the particular mode in question throughout the material. A good approximation to a single mode resonator can however be obtained by the proper choice of geometry and dimensions.

In the presence of an harmonic driving term of a given frequency, the onset of vibrations will be accompanied by the excitation of normal modes. In a real, lossy, material these form transients dying away exponentially, leaving a steady-state vibration of the same frequency as the driving force. The amplitude of the forced vibrations will generally be of the same order of magnitude as the static deformation produced by a force equal to the amplitude of the periodic driving force, except when the driving frequency is close to the frequency of a normal mode of the device, when resonance

occurs. At resonance, the amplitude of vibration increases by an amount limited only by dissipation, in analogy to the increase in current observed when the frequency of an applied voltage is swept through the resonance frequency of a series tuned *LCR* circuit. The analogy is almost exact for frequencies close to the resonance frequency, but as the electrical circuit has only the one resonance frequency, it does not reflect the multiplicity of resonance frequencies that exist in the mechanical vibrator.

In resonant systems where the resonator material has low losses and the driving force is only loosely coupled to the actual resonator, the driving term can be largely ignored in respect of the calculation of resonance frequencies, simply being regarded as a means of maintaining the vibration. Both these conditions are true in the case of quartz resonators, thus simplifying the analysis of the device by allowing attention to be restricted to free vibrations only.

QUARTZ RESONATORS

A piezoelectric resonator consists of a piece of piezoelectric material precisely dimensioned and oriented with respect to the crystallographic axes of the material and equipped with one or more pairs of conducting electrodes. By means of the piezoelectric effect an electric field applied between the electrodes excites the resonator into mechanical vibration. The amplitude of vibration is negligibly small except when the frequency of the driving field is in the vicinity of one of the resonator's normal modes of vibration and resonance occurs. Near resonance, the amplitude of vibration increases and, again by virtue of the piezoelectric effect, the electrical impedance of the device changes rapidly.

Many different substances have been investigated as possible resonator materials, but for many years quartz resonators have been preferred in satisfying needs for precise frequency control and selection. Compared to other resonators, for example *LC* circuits, mechanical resonators such as tuning forks, and piezoelectric resonators based on ceramics or other single-crystal materials, the quartz resonator has a unique combination of properties. The material properties of single-crystal quartz are both extremely stable and highly repeatable from one specimen to another. The acoustic loss or internal friction of quartz is particularly low, leading directly to one of the key properties of a quartz resonator, its extremely high *Q* factor. The intrinsic *Q* of quartz is 10^7 at 1 MHz. Mounted resonators typically have *Q* factors ranging from tens of thousands to several hundred thousand, orders of magnitude better than the best *LC* circuits.

The second key property of the quartz resonator is its stability with respect

to temperature variation. Depending on the shape and orientation of the crystal blank, many different modes of vibration can be used and it is possible to control the frequency-temperature characteristics of the resonator to within close limits by an appropriate choice. The most commonly used type of resonator is the AT-cut where the quartz blank is in the form of a thin plate cut at an angle of about $35^{\circ} 15'$ to the optic axis of the crystal. The AT-cut has a frequency-temperature coefficient described by a cubic function of temperature, which can be precisely controlled by small variations in the angle of cut. This cubic behaviour is in contrast to most other crystal cuts which show parabolic temperature characteristics, and makes the AT-cut well suited to applications requiring a high degree of frequency stability over wide temperature ranges.

The third essential characteristic of the quartz resonator is related to the stability of its mechanical properties. Short and long term stabilities manifested in frequency drifts of only a few parts per million per year are readily available even from commercial units. Precision crystal units manufactured under closely controlled conditions are second only to atomic clocks in the frequency stability and precision achieved.

The AT-cut resonator uses the thickness shear mode of vibration. A standing wave is set up in the crystal blank by the reflection at both major surfaces of transverse waves travelling in the thickness direction. The major mechanical displacement is in the plane of the crystal, at right angles to the direction of wave propagation. At resonance, an odd number of half wavelengths is contained in the thickness of the plate, the thickness therefore being the primary frequency determining dimension. The AT-cut is thus an example of a thickness mode resonator. AT-cuts are commonly manufactured in the frequency range from about 1 MHz to 200 MHz and above, and in this range are usually the optimum choice for most applications. The AT-cut is, however, sensitive to stresses in the body of the resonator, whether caused by temperature gradients due to rapid external temperature changes, or by external forces. For applications where extreme stability is required, or where severe environmental conditions are likely to be encountered, this stress sensitivity is a drawback, and newer crystal cuts such as the SC have been developed that minimize these effects. The SC-cut is also a thickness mode resonator, and as such is basically available in much the same frequency range as the AT, though its commercial availability is presently much more limited.

Below about 1 MHz, thickness-mode resonators generally become too cumbersome and unwieldy for general use, and other modes of vibration are used. At very low frequencies, below about 100 kHz, flexural mode and length extensional devices are used. In both cases the crystal is cut in the form of a long thin bar with the length of the bar being the primary frequency determining dimension. Above 100 kHz face shear devices, such as the CT- and DT-cuts, are used, where the crystal element is a square or rectangular

plate, and the length of the edges determines the frequency. The GT-cut also uses a rectangular plate vibrating in coupled length and width extensional modes, and is unique among crystal cuts in having an almost zero frequency-temperature coefficient over a wide range of temperature. The traditional method of manufacture for all these low frequency crystal cuts involved mounting the crystal element by means of thin wires attached to the crystal at nodal points and soldered or welded to the crystal holder. Such wire-mounted units are intrinsically less robust than the higher frequency thickness mode resonators. Recent developments triggered by the huge demand for miniature crystal units for electronic watches have however initiated a renaissance in the technology of low frequency crystal manufacture, typified by the introduction of a miniature GT-cut crystal in which both resonator and support structure are fabricated as one piece using photolithographic techniques. These techniques, originally developed for the mass production of watch crystals of the tuning fork type using flexural vibrations, are now also being applied to the miniaturization of AT type resonators.

CRYSTAL OSCILLATORS AND FILTERS

Crystal filters and oscillators are devices in which one or more resonant circuits are realized by crystal resonators in order to benefit from their superior Q , and their frequency stability with respect to temperature and time. In the case of filters, the high Q available from quartz resonators in particular is essential in the practical realization of effective narrowband filters, and crystal filters have long been a key component in frequency division multiplex (FDM) line transmission systems. In modern radio communications practice, crystal filters play an equally essential role in determining the selectivity of radio systems. The latter role has largely developed since the 1950s and was greatly accelerated by the introduction of the monolithic crystal filter in the 1960s. In the monolithic filter, not just one but two or more coupled resonant circuits are realized on a single piece of quartz, resulting in large cost and space savings and much increasing the attractiveness of crystal filters to the equipment designer.

In the early years following Cady's pioneering work on the crystal controlled oscillator, the only applications envisaged were in frequency standards and other purely scientific areas. The need for more general crystal control of oscillators was not recognized until 1926, when the problems of poor quality reception being experienced by listeners to the radio station WEAf in New York were solved by the introduction of crystal control of the transmitted signal (Heising, 1946). Subsequent developments in both communications and crystal technology have brought about a situation where

crystal control is vital to the present intensive use of the radio frequency spectrum.

For much of the period from Cady's work to the present, the design of the crystal oscillator was the province of the user, the equipment manufacturer. The crystal manufacturer supplied only the crystal. Although there have always been exceptions to this general rule, the growth of a volume market for packaged crystal oscillators is a relatively recent phenomenon, driven by several factors including the impact of frequency synthesis techniques, the increasing need for temperature compensated devices, and latterly the scarcity of design engineers familiar with analog circuits. In non-synthesized multi-channel radio equipments, the normal design approach is to obtain the desired channel frequencies by switching appropriate crystals, and there is no scope for packaged oscillators. However, in a synthesized equipment only one reference oscillator is required from which all necessary channel frequencies are derived, and it is clearly advantageous for the equipment designer to be able to specify the reference oscillator as a bought-in component rather than have to go through a complete design exercise. This is especially true when the stability requirements are such that temperature compensation of the oscillators is required. It is a feature of the temperature compensation problem that whatever method of compensation is adopted has to be tailored to individual crystals. This means that in a non-synthesized system all channel crystals have to be separately compensated, whereas in a synthesized system only the reference frequency needs compensation. Again from the equipment designer's viewpoint, it is surely preferable to be able to specify the stability required in a packaged oscillator and not to have to worry about the niceties of temperature compensation.

The emphasis in modern electronics is very much on digital techniques, which are increasingly being used in areas such as communications previously dominated by analog methods. For an engineer with a digital electronics background, the design of a crystal controlled oscillator is likely to present problems that can be avoided by the specification and procurement of a packaged oscillator. This is particularly relevant in the case of wholly digital systems where an oscillator is required simply to produce timing signals to allow the synchronization of different digital processes.

As a result of these various factors a wide range of packaged crystal oscillators is now commercially available. At one extreme are the 'clock' oscillators designed for just such timing applications as mentioned above. These are manufactured in large volumes in standard dual-in-line configurations. Their frequency stability is relatively poor, but adequate for the application, and they are of course designed to interface directly with digital circuits. At the other extreme are crystal oscillators incorporating miniature crystal ovens and achieving frequency stabilities of parts in 10^8 over wide temperature ranges. In between these extremes lies a large variety of different oscillator types, including, for example, voltage controlled oscillators

(VCXOs) and temperature compensated crystal oscillators (TCXOs), which allows an equipment engineer to concentrate on the larger issues of system design rather than the details of crystal oscillator design.

In crystal filters, too, there is currently available a wide range of devices for performing narrowband filtering functions. Apart from line transmission applications, the bulk of filters sold are IF or 'channel' filters for VHF and UHF radio systems. Other filter types include SSB filters for eliminating unwanted sidebands, 'roofing' filters for double-conversion HF receivers, narrowband bandstop filters, and various filters with specialized phase and group delay responses. In the channel filter sector there is a good deal of standardization, but in other sectors the crystal filter market is still primarily one for custom-built devices.