PULSED NONLINEAR SURFACE ACOUSTIC WAVES IN CRYSTALS

R. E. Kumon, M. F. Hamilton, Yu. A. Il'inskii, E. A. Zabolotskaya, P. Hess, A. M. Lomonosov, and V. G. Mikhalevich

16th International Congress on Acoustics and 135th Meeting of the Acoustical Society of America, Seattle, WA, USA, 20–26 Jun 1998

SLIDE NOTES

Notes on COVER PAGE

- The experimental work was supported by Volkswagen-Stiftung, Deutche Forschungsgemeinschaft, and the Russian Foundation for Basic Research.
- The theoretical work was supported by the U.S. Office of Naval Research.

- Yu. A. Il'inskii and E. A. Zabolotskaya are now employed by MacroSonix Corporation, Richmond, Virginia, USA.

Notes on ANISOTROPY IN CRYSTALLINE SILICON

- This talk will focus on surface acoustic waves (SAW) in crystalline silicon. Crystalline silicon is a cubic crystal (see "Diamond Cubic Structure" diagram). Cubic crystals have the stress-strain relation shown in the slide. Usually the strains are sufficient small that the linear relation $\sigma_{ij} = c_{ijkl} e_{kl}$ is valid. However, the strains considered here are large enough that the nonlinear terms contribute significantly and, in fact, give rise to shock formation.
  - Diamond Cubic Structure: Si has a diamond lattice which can also be considered to be two fcc lattices, one displaced relative to the other by $(1/4, 1/4, 1/4)$. Note also that every atom has four nearest neighbors. The lattice spacing for Si is 0.543 nm.
- The material properties of the crystal are expressed via the elastic constants of the material. In particular, 3 SOE and 6 TOE constants are necessary to specify a cubic crystal. For all the simulations presented in this talk, the data for the Si elastic constants was taken from the paper by H. J. McSkimin and P. Andreatch, Jr., *J. Appl. Phys.* **35**, 3312–3319 (1964).
- Because these systems are anisotropic, the wave propagation is different depending on how the crystal is cut and the direction that the wave is travelling.
  - The surfaces of cut crystals have traditionally been described using a crystallographic convention called Miller indices. Miller indices are defined by finding three noncollinear atoms on the surface that intersect the crystal axes and then applying the following method:
    1. Find the intercepts of the three basis axes in terms of the lattice constants.
    2. Take the reciprocals of these numbers and reduce to the smallest three integers having the same ratio. The result is enclosed in parentheses (hkl). [from C. Kittel, *Introduction to Solid State Physics*, 2nd ed. (John Wiley & Sons, New York, 1965), p. 34] Note that if the Miller indices are interpreted as a vector components, the resulting vector is normal to the surface of the cut.
  - Directions are specified in a different way:

    The indices of a direction in a crystal are expressed as the set of the smallest integers which have the same ratios as the components of a vector in the desired direction referred to the axis vectors. The integers are written in square brackets, [uvw]. The x axis is the [100] direction; the $-$y axis is the $[0\bar{1}0]$ direction. A full set of equivalent directions is denoted this way: $\langle$uvw$\rangle$. [from C. Kittel, *Introduction to Solid State Physics*, 2nd ed. (John Wiley & Sons, New York, 1965), p. 34]

    This presentation will use both of these notations frequently.
- The crystal cut chosen in the experiment is the (111) plane (see diagram). For the experiment and simulations shown in this talk, the pulses propagated in the $\langle 112 \rangle$ direction.

Notes on SCHEMATIC DIAGRAM: SURFACE ACOUSTIC WAVE

- This figure shows the particle motion of a typical surface acoustic wave in an anisotropic medium. Consider the case of a surface acoustic wave with an initially sinusoidal velocity waveform in an isotropic and anisotropic material. Assume that the x-axis is in the propagation direction and that the z-axis is normal to the surface cut.
- In the most general case, the particle motion is elliptical with the plane of the ellipse rotated by some angle ϕ out of the xz-plane. Different perspectives of this transverse motion are shown in the top and front views.
- Due to the symmetry of the particular cut and direction considered here, $\phi = 0$. Hence there is no transverse component to the particle motion. Note, however, that further investigation has shown that transverse motion does occur for other directions in the same plane. This situation differs from the isotropic case in that transverse motion does not exist for waves propagating in any direction.

Notes on THEORY

- Briefly, the approach used here involves calculating the Hamiltonian energy function through cubic order in the wave variables, choosing appropriate generalized coordinates, applying the equations of motion in canonical form, and deriving evolution equations for the slowly varying amplitudes in a suitable retarded time frame. The approach was outlined in M. F. Hamilton, Yu. A. Il'inskii, and E. A. Zabolotskaya, "Nonlinear surface wave propagation in crystals," *Nonlinear Acoustics in Perspective*, R. J. Wei, ed. (Nanjing University Press, Nanjing, China, 1996), pp. 64–69.
  - Note that computing the Hamiltonian the quadratic order would only give rise to linear terms in the model equations. Thus, the potential energy terms to at least cubic order in the strain must be included to model nonlinear effects.
  - Note also that this method is very general. It is applicable to any elastic material for which the SOE and TOE constants are known and to any cut and direction in such a material.
- Assumptions:
  1. It is assumed that the nonlinear solution is close to the linear solution; in particular the depth dependence of each frequency is the same as in the linear solution.
  2. It is assumed that the wave fronts are planar.
  3. It is assumed that the wave is progressive, i.e., travels only in one direction. (It should be possible to extend the theory to include compound waves; only the results will be more complicated.)
- The velocity waveforms in the solid are assumed to take the form shown in the slide.
  - The coordinate system for the solution is chosen such that the the z-axis is perpendicular to the surface of the solid and the x-axis is in the direction of the propagation of the wave.
  - Note that on the surface they simplify to

$$v_i(x,z,t) = \sum_{n=-\infty}^{\infty} v_n(x) \sum_{s=1}^{3} \beta_i^{(s)} e^{in\tau} \qquad [v_n^* = v_{-n}]$$

    where $\tau = k_0 x - \omega_0 t$ is the retarded time and the $\beta_i^{(s)}$ are constants determined from the linear problem.
  - Note that surface acoustic waves are non-dispersive, i.e., their wave speed is not frequency dependent.
- The coupled, nonlinear spectral evolution equations that result from this approach are shown above. Here v_n is the complex amplitude of the nth harmonic, α_n is the attenuation coefficient for the nth harmonic, and R_{lm} is the nonlinearity matrix.
  - The *ad hoc* attenuation term $\alpha_n = n^2 \alpha_1$ is added to the left-hand side for purposes of numerical stability when solving the equations. For all the cases shown here the dimensionless value of $\alpha_1 = 0.025$. This attenuation is sufficiently weak that its main effect is to stabilize the portion of the waveform in the neighborhood of the shock without significantly the remainder of the waveform. Note that the dimensionless value of α_1 here is the analog of the Goldberg number Γ for nonlinear acoustic waves in fluids.

  - Physically, the nonlinearity coefficients R_{mn} represent the strength of the coupling between different harmonics in the wave. They are given by a complicated analytical expression which can be determined completely by knowing the SOE and TOE constants of the material. (See the supplementary notes below for the equation for R_{mn} and more information.)
  - For the case of isotropic materials, these equations can be shown to reduce to the evolution equations previously derived by Zabolotskaya [E. A. Zabolotskaya, "Nonlinear propagation of plane and circular waves in isotropic solids," *J. Acoust. Soc. Am.* **91**, 2569–2575 (1992)].
  - While Hamilton's equations describe the evolution of a system in time, the evolution equations listed in the slide evolve in space, not time. Informally speaking, the transformation between the two is done by moving into retarded time frame and thereby replacing $\partial/\partial t$ with $c\,\partial/\partial x$. It is possible to demonstrate formally that this is the proper transformation and that it is not an approximation [E. Yu. Knight, M. F. Hamilton, Yu. A. Il'inskii, and E. A. Zabolotskaya, "General theory for the spectral evolution of nonlinear Rayleigh waves," *J. Acoust. Soc. Am.*, **102**, 1402-1417 (1997)].

- The spectral evolution equations were solved numerically. First the nonlinearity matrix elelments R_{lm} were computed from the material constants, and the spectral "source" condition was computed from the data measured at $x = 5$ mm from the excitation point (see diagram on the "Experiment" slide for the experimental setup). Next, a fourth-order Runge-Kutta routine was used to integrate the system. The waveform expansions used had 400 harmonics. With a pulse repetition frequency of 10 MHz, the bandwidth realizable by the simulations was then 4000 MHz.
  - In theory, there are an infinite number of equations to integrate. For purposes of computation, the velocity waveform expansions were truncated such that only terms with $n = -400$ to $n = 400$ were included in the sum. However, because the velocity waveforms must be real-valued, $v_{-n} = v_n^*$. Therefore only 400 spectral amplitudes must be determined and, correspondingly, only 400 equations must be integrated.
  - Strictly speaking, the Fourier series expansion of a function can only be done if the function is periodic. To allow the pulse to be modelled with a relatively small number of harmonics, it is necessary then to assume that the "pulse" is not a single event but one in a series that repeats periodically. For these simulations, the pulsed signal was assumed to repeat every 100 ns, thereby giving a pulse repetition frequency of 10 MHz. This is then the lowest frequency component that can be modelled by the simulation. With 400 harmonics, this also makes 4000 MHz the highest frequency that can be modelled.

Notes on EXPERIMENT

- The approach used here generates SAW via thermoelastic laser excitation. This method was described previously by A. Lomonosov and P. Hess, "Laser excitation and propagation of nonlinear surface acoustic wave pulses," *Nonlinear Acoustics in Perspective*, R. J. Wei, ed., (Nanjing University Press, Nanjing, China, 1996), pp. 106–111.

- The basic setup is shown in the diagram. The SAW pulse was generated by a Nd:YAG laser that was focused with a cylindrical lens into a thin strip 6 mm by 50 μm on the surface of crystal. To detect the resulting SAW pulse, optical probe beams were employed. This can be done because probe beam deflection is proportional to the vertical velocity component v_z at the surface. The probe beam deflections were detected by split photodiodes with a bandwidth of 500 MHz. The probe beams irradiated spots approximately 4 μm in diameter on the surface at distances 5 mm and 21 mm from the excitation region.
  - Experimental details:
    The Nd:YAG laser that generated the surface waves had a wavelength of 1064 nm (infrared), pulse duration 7 ns, and energy up to 50 mJ. A strongly absorbing carbon layer in the form of an aqueous suspension was placed on the surface of the Si in the excitation region to facilitate energy transfer to the surface and prevent cracking. The transient SAW waveforms were then measured absolutely with a calibrated probe-beam deflection setup using stabilized cw Nd:YAG laser probe beams with a wavelength of 532 nm (green visible) and power of 40 mW.

- Surface wave pulses in these experiments had durations of 20–40 ns and peak strains between 0.005 and 0.010. As will be shown below, pulses of this magnitude exhibit nonlinear behavior that give rise to waveform distortion and shock formation.

- For example, a previous experiment (described in the Lomonosov and Hess reference above) generated SAW in Si with particle velocities of over 40 m/s. Because the surface acoustic wave speed for silicon in this direction is 4730 m/s, the Mach number equals 0.0085. To compare, the equivalent Mach number in air at room temperature is equivalent to a SPL of 155 dB (re 20 μPa). The same experiment generated peak-to-peak particle displacements of over 700 nm. Because the lattice constant for Si is only 0.543 nm, this means that the displacement is nearly 1300 times the size of the atoms in the crystal.
- In fact, the strains are so large that in some cases the pulses generated caused the Si crystal to fracture. The optical absorptive liquid placed on the surface (described above) helps to prevent this from occurring.

Notes on DIFFRACTION EFFECTS

- Because the theory presented above models only plane waves, it is necessary to determine if the effects of diffraction are significant in the SAW beam that is produced by the experiment.
- To do this, first examine the measured frequency spectrum shown in the diagram at the "source" probe location of $x = 5$ mm. The peak in the spectrum occurs around 50 MHz.
- With this information, the effects of diffraction can be estimated by computing the diffraction length of the beam. With a characteristic beam radius of 3 mm and the characteristic frequency of 50 MHz, the diffraction length for the SAW beam is computed to be around 300 mm. However, the furthest measurement distance here is only 21 mm, a full order of magnitude less than the diffraction length.
  - The diffraction length, as known as the Rayleigh distance, is the approximate distance past which the wavefronts in the beam are no longer essentially planar. It marks the transition between the near-field and far-field regions of the beam. For more information on nonlinear sound beams, see M. F. Hamilton, "Sound beams," Chap. 8 in *Nonlinear Acoustics*, M. F. Hamilton and D. T. Blackstock,, eds. (Academic Press, Boston, 1998), pp. 233–261.
- The conclusion here is then that diffraction effects are not important, and the theoretical predictions should be valid in this regime.

Notes on SIMULATIONS OF SINUSOIDS

- This slide shows simulations performed with a monofrequency, continuous wave, sinusoidal signal as the source condition ($x = 0$ mm) at several distances. As shown here, the vertical velocity is a cosine wave while the longitudinal velocity is a sine wave.
- The top diagram contains the longitudinal velocity waveform while the bottom figure contains the vertical velocity waveform. Due to the symmetry of this particular case, there is no transverse velocity component. The longitudinal velocity is computed from vertical velocity by linear theory. The waveforms were reconstructed with a bandwidth of 4000 MHz to match the maximum bandwidth of the numerical simulations.
- The period of the wave was taken to be 20 ns so that the characteristic frequency is 50 MHz as in the pulse data. In addition, the amplitude of the vertical velocity component was selected to be 37 m/s, close to the characteristic velocity (half the peak-to-peak velocity) of the pulse data.
- From these characteristic values, an estimate of the shock formation distance can then be computed from theory to be around 2.9 mm. Because the pulse propagation distance $\Delta x = 16$ mm, the pulse is well into the shock formation region.
  - The estimate of the shock formation distance is computed via a quasilinear analysis of the harmonic generation. See D. J. Shull, M. F. Hamilton, Yu. A. Il'insky, and E. A .Zabolotskaya, "Harmonic generation in plane and cylindrical nonlinear Rayleigh waves," *J. Acoust. Soc. Am.* **94**, 418–427 (1993) for an example of how this is done for isotropic materials.
- Notice also that the waveform distortion is asymmetric and the peaks increase significantly in magnitude. Going forward in time from the trough to the peak, the vertical velocity waveform steepens, while going from the peak to trough, the waveform flattens. In the longitudinal velocity waveform, the peaks recede in time while the troughs advance in time. This behavior, which is opposite to the distortion of a sound wave in a fluid, occurs because the coefficient of nonlinearity is negative in this case. Similar distortions will be seen in the pulse data.

  - The cusped peaks increase in magnitude because as the waveform distorts more of its energy is shifted into higher frequency components. Because the penetration depth of a SAW is proportional to its wavelength, these higher frequency and lower wavelength SAW have more of their energy near the surface. It is this effect that causes the large peaks. Corresponding snapshots of the SAW further beneath the surface would show that the peaks of the waveforms at those depths would decrease in magnitude so that energy is conserved.

Notes on WAVEFORMS AT FIRST LOCATION

- This slide shows the experimental data at the first probe beam location ($x = 5$ mm).
- Notice that the vertical velocity waveform has the general trends of the cosine signal shown in the vertical velocity waveform on the previous slide. Hence it may be expected that the portion of the waveform going forward in time from the trough to the peak will steepen, while the portion going from the peak to the trough will flatten.
- Similarly, the latter part of the longitudinal velocity waveform has the general trend of the sine signal shown in the previous slide. Hence it may be expected that the larger peak will recede and the trough will advance. The smaller peak will also recede but less than the larger peak. The net result will be that the the pulse will lengthen in time.

Notes on EVOLUTION OF WAVEFORMS

- This slide shows the evolution of the experimental waveforms from the first probe beam location ($x = 5$ mm) to the second probe beam location ($x = 21$ mm). The experimental (solid line) and theoretical (dashed line) waveforms are also compared at the second beam location.
- The comparison shows that the theory is in close quantitative agreement with the experiment, and the waveforms evolve generally as expected based upon the discussion of the sinusoidal waveform simulations above. Here the experimental waveforms have a bandwidth of 500 MHz (from the photodiodes) while the theoretical waveforms were reconstructed with bandwidth of 700 MHz. Note the distinct lengthening of the pulse as seen most clearly by comparing the longitudinal velocity waveforms (it can also be seen in the supplementary notes below).
- One aspect of the waveforms that does not seem to match the sinusoidal simulations is that the pulse height does not increase significantly. As will be shown next, this is probably due to the fact that the bandwidth limitations of the photodiodes exclude the higher frequency terms necessary to record steep shocks and narrow peaks. In other words, the temporal resolution of the experiment (1 ns) is probably insufficient to resolve the very short peaks in the evolved waveforms.

Notes on EFFECT OF RECONSTRUCTION BANDWIDTH

- This slide shows the experimental data for the longitudinal velocity at the second probe location ($x = 21$ mm) along with two theoretical waveforms reconstructed with different bandwidths. In both of the theoretical waveforms, the spectral amplitudes were computed out to 4000 MHz, but the amplitudes where weighted by the shading function shown in the diagrams on the right side of the page.
  - While it is possible to reconstruct the waveforms with all 4000 MHz of spectral data, this actually gives less accurate results because the highest frequencies contain the energy aliased from all even higher frequencies. This is purely a numerical effect that arises from truncating the frequency spectrum at a finite value.
- The top set of figures shows the theoretical waveform reconstructed with 700 MHz of shaded spectral data. In other words, each spectral amplitude was multiplied by the weighting function $\exp(-[f/700])$ before reconstruction in the time domain (here f is the frequency of the spectral amplitude) . The frequency content of the resulting theoretical waveform is then close to the frequency content of the experimental waveform (500 MHz), and the peak height matches closely.
- The bottom set of figures shows the theoretical waveform reconstructed with 3000 MHz of the spectral data. Because the frequency content of the resulting theoretical waveform is significantly greater than the frequency content of the experimental waveform (500 MHz), the theoretical waveform can then resolve sharper peaks. These large peaks then parallel the large peaks seen above in the sinusoidal simulations.

- Hence it appears likely that if the bandwidth of the photodiodes of the experiment were larger, then sharper peaks would be observed.

Notes on CONCLUSION & FUTURE WORK

- To conclude, this is the first reported comparison of experiment and theory for nonlinear SAW in a crystal. The theory is in close quantitative agreement with the experiment and the predictions of the theory are based on the fundamental properties of the material (density, SOE, TOE).
- Future work will include studying the relationship between the nonlinearity matrix elements and waveform distortion (to be discussed at the Norfolk ASA meeting), studying the variation of waveform evolution as function of cut and direction, investigating other anisotropic materials, and investigating piezoelectric effects.

Supplement: NONLINEARITY MATRIX & LINEAR THEORY

The nonlinearity matrix is given by

$$R_{n_1 n_2} = -\sum_{s_1,s_2,s_3=1}^{3} \frac{d'_{iklmpq}\beta_i^{(s_1)}\beta_l^{(s_2)}\beta_p^{(s_3)*}l_k^{(s_1)}l_m^{(s_2)}l_q^{(s_3)*}}{2[n_1 l_3^{(s_1)} + n_2 l_3^{(s_2)} + (n_1+n_2)l_3^{(s_3)*}]}$$

where $\beta_i^{(s)} = C_s\alpha_i^{(s)}$. To compute this expression, the linear problem must first be solved.

Start with linearized wave equation

$$\rho\frac{\partial^2 u_i}{\partial t^2} = \frac{\partial \sigma_{ij}}{\partial x_j} = c_{ijkl}\frac{\partial^2 u_k}{\partial x_j \partial x_l} \ . \tag{1}$$

Next assume SAW solution of form

$$u_i = \sum_{s=1}^{3} C_s\alpha_i^{(s)} e^{ik(\mathbf{l_s}\cdot\mathbf{r}-\omega t)} \tag{2}$$

where $l_s = \{1, 0, \zeta\}$. Substitute Eq. (2) into Eq. (1) to yield

$$\rho c^2\alpha_i = \tilde{c}_{ijkl}l_j l_l \alpha_k \ . \tag{3}$$

Solve Eq. (3) subject to the stress-free surface boundary condition

$$\sigma_{i3} \mid_{x_3=0} = 0 \ . \tag{4}$$

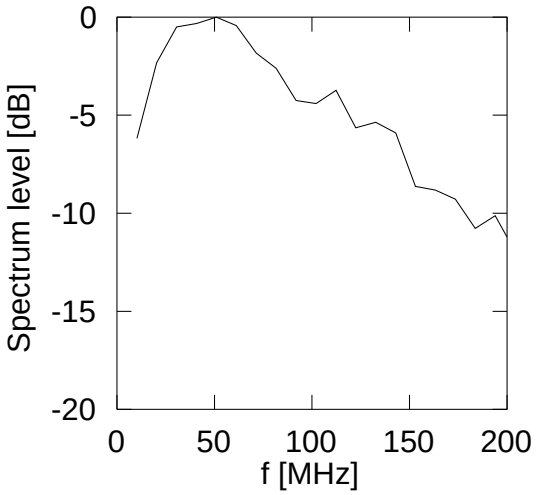
Substituting Eq. (2) into Eq. (4) yields

$$\tilde{c}_{i3kl}\sum_{s=1}^{3} C_s\alpha_k^{(s)}(c)l_l^{(s)} = 0 \ . \tag{5}$$

This equation can be solved numerically.

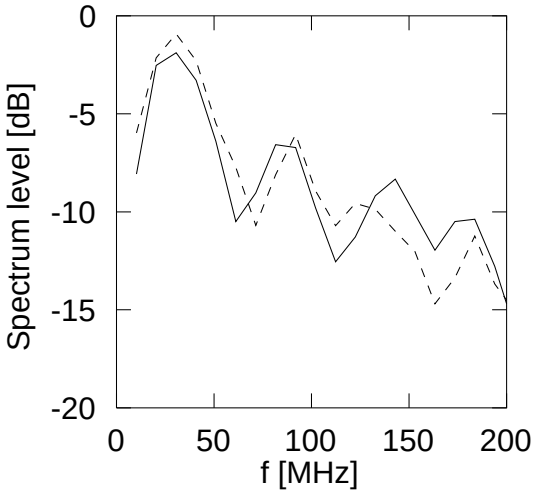
Supplement: EVOLUTION OF SPECTRA

- The experimental spectrum at $x = 5$ mm is shown below:



The spectrum is normalized such that the peak spectral amplitude is unity.

- The experimental spectrum (solid line) and theoretical spectrum (dashed line) at $x = 21$ mm is shown below:



This spectrum is normalized relative to the peak spectral amplitude of the previous figure. The effects of dissipation can be seen across the entire spectrum. In addition, harmonic generation is evident from the multiple peaks of the spectrum. This spectrum also clearly shows the effect of pulse lengthening as the peak frequency shifts downward from about 50 MHz to about 33 MHz.