Dynamical Model for the Absorption and Scattering of Ballistic Phonons by the Electron Inversion Layer in Silicon

J. C. Hensel
Bell Laboratories, Murray Hill, New Jersey 07974

and

B. I. Halperin
Physics Department, Harvard University, Cambridge, Massachusetts 02138

and

R. C. Dynes
Bell Laboratories, Murray Hill, New Jersey 07974

(Received 4 August 1983)

A new, dynamical model is presented which can explain the anomalously strong attenuation reported for ballistic phonons propagating through the (001) inversion layer of Si. The crux is an interference effect between phonon amplitudes backscattered from the two-dimensional electron gas and amplitudes specularly reflected from the Si-SiO_2 interface.

PACS numbers: 63.20.Kr, 66.70.+f, 73.40.Qv

Recently we reported experiments[1,2] employing ballistic phonons propagating through the (001) inversion layer of Si as a means of measuring the electron-phonon interaction in a two-dimensional electron gas. Surprisingly, the observed attenuation of phonon intensity (ranging from 1% to 2%) was an order of magnitude too large to be explained by absorption. This puzzling result is all the more fascinating in view of persistent suspicions that the electron-phonon interaction might be much larger in two dimensions (2D) than in 3D.[3] In this Letter we develop a new theory of phonon scattering and absorption for the inversion layer which successfully accounts for these observations in terms of the known electron-phonon coupling in 3D and dynamical properties of a 2D electron gas.

The experiments were conducted in a reflection scheme (see Fig. 1); so in our original interpretation[1,2] we used the simple model of absorption of ballistic phonons passing twice through the inversion layer, before and after specularly reflecting from the Si-SiO_2 interface. If the reflection were total, this might not be a bad approximation. But in the present case the reflectance is small[4] ($r \sim 4\%$) as θ is close to the Brewster angle, and as we shall see below the largest contribution is an interference between the phonon amplitude backscattered from the inversion layer and the reflected primary beam. This term is enhanced by roughly a factor of $r^{-1/2} \sim 5$.

Our approach is to calculate the linear response of the 2D electron gas to perturbations by the phonons. The radiation fields thus set up are contributed by both real and imaginary parts of the complex density response function $\chi(\vec{q}, \omega)$ which relates the 2D electron density $n_2(x, y)$ to a perturbing potential (at wave vector $\vec{q}$ in the plane and frequency ω) which couples linearly to $n_2(x, y)$. This is a dynamical theory whereas the earlier one was "static" in the sense that it dealt exclusively with $\mathrm{Im}\chi(\vec{q}, \omega)$ as an absorption mechanism.

Our first concern is to formulate the coupling between the phonons and the 2D electron gas. We shall regard the phonons classically as elastic waves; the quantization plays no role whatsoever

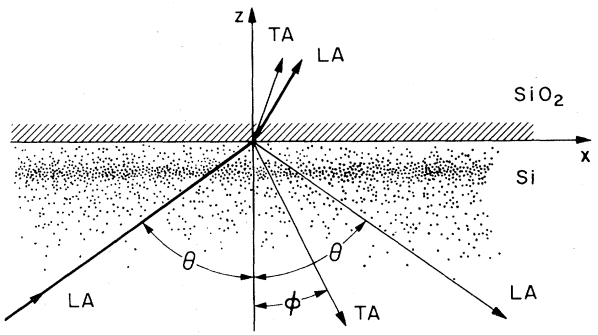

FIG. 1. Schematic representation of the experiment (Refs. 1 and 2) showing the interface, the electron inversion layer (shading), and the incident, reflected, and transmitted phonons (LA and TA denote longitudinal and transverse acoustic modes, respectively). The LA intensity is largely transmitted as signified by heavy lines.

in the present situation. We adopt a coordinate system shown in Fig. 1 in which z is normal to the Si-SiO_2 interface (the oxide occupies the half space $z>0$). It will be assumed that Si and SiO_2 are both elastically isotropic. The coupling between phonons and electrons in the occupied [001] valleys is given by an interaction Hamiltonian containing the deformation potential[5] which acts on the electron density $n(\vec{r})$,

$$\mathcal{H}' = \int d^3r\, n(\vec{r})[a(\epsilon_{xx}+\epsilon_{yy})+b\epsilon_{zz}]. \quad (1)$$

Here a and b are deformation-potential constants[6] and ϵ_{ij} is the strain tensor $2^{-1}(\partial u_i/\partial x_j+\partial u_j/\partial x_i)$, $\vec{u}$ being the displacement. There is no loss of generality if we let the phonon wave vectors be in the xz plane, whereupon $\epsilon_{yy}=0$. Next, we assume that $n(\vec{r})=n_2(x,y)f(z)$, $f(z)$ being the electron probability density normal to the xy plane, i.e., $f(z)=\zeta_0^2(z)$ where $\zeta_0(z)$ is the wave function of the lowest subband 0. Then by definition the response to phonons can be expressed as a Fourier transform in the xy plane,

$$n(\vec{q},z) = -f(z)\chi(\vec{q},\omega)\int dz' f(z')[a\epsilon_{xx}(\vec{q},z')+b\epsilon_{zz}(\vec{q},z')], \quad (2)$$

where ω is the phonon frequency and $\vec{q}$ is the projection of the phonon wave vector in this plane.

The equations of motion for the displacement $\vec{u}$ are

$$\rho\omega^2\vec{u}+\nabla\cdot\vec{\sigma}=\begin{cases}0, & z>0\\ \delta\mathcal{H}'/\delta\vec{u}, & z<0\end{cases} \quad (3)$$

where ρ is the crystal density, $\vec{\sigma}$ is the stress tensor related to $\vec{\epsilon}$ in the usual way through the elastic moduli (the elastic moduli and density are different in the two regions $z>0$ and $z<0$), and the functional derivative $\delta\mathcal{H}'/\delta\vec{u}$ represents the driving force due to the electrons, which from (1) can be written

$$\frac{\delta\mathcal{H}'}{\delta u_x}=-a\frac{\partial n(\vec{r})}{\partial x}, \quad \frac{\delta\mathcal{H}'}{\delta u_z}=-b\frac{\partial n(\vec{r})}{\partial z}. \quad (4)$$

Solutions to (3) in the oxide (source-free) half space are trivial (i.e., plane waves propagating in the $+z$ direction); but in the half space $z<0$ occupied by the electron gas, which acts as a distributed source, they are not and in general necessitate numerical methods. We have elected instead to employ a simpler "slab" model wherein the density is replaced by a rectangular profile of thickness s and spaced a distance h from the interface at $z=0$, i.e., $f(z)=s^{-1}[\theta(z+h)-\theta(z+h-s)]$, where $\theta(z)$ is the unit step function. Now one finds solutions of the form

$$\vec{u}=[A_l\vec{\epsilon}_l\exp(ik_lz)+A_j\vec{\epsilon}_j\exp(ik_tz)+A_m\vec{\epsilon}_m(\exp -ik_lz)+A_n\vec{\epsilon}_n\exp(-ik_tz)]e^{iqx} \quad (5)$$

in each of the four regions $z>0$, $0>z>-h+s$, $-h+s>z>-h$, and $-h>z$, with appropriate matching conditions at the boundaries. The interface at $z=0$ one treats as a standard Fresnel problem for elastically dissimilar media with the usual boundary conditions of continuity of σ_{zx}, σ_{zz}, u_x, and u_z. At the profile edges, $z=-h$ and $z=-h+s$, the displacements u_x and u_z are continuous, but the stress tensor is not. One sees from (3) and (4) that the discontinuities in $f(z)$ require discontinuities in σ_{zz}, the magnitudes of which, $bn_2(\vec{q})/s$, follow from integration of (3) across the boundaries. Similarly, the electron density gives a contribution to σ_{zx} approximated by a discontinuity $iaqn_2(\vec{q})/2$ at each profile edge. These discontinuities represent sources of elastic radiation due to oscillations in electron density. The A_i in (5) are field amplitudes, $\vec{\epsilon}_i$ are unit polarization vectors, and k_l and k_t are, respectively, the normal components of the LA and TA wave vectors.

Imposition of the aforementioned conditions at the three boundaries results in a system of twelve linear equations for the twelve (generally complex) amplitudes $A_3,\ldots,A_{14}$ (A_1 and A_2 are the given incident LA and TA amplitudes, respectively) to be solved simultaneously with Eq. (2) [if we regard $n_2(\vec{q})$ as an "unknown" and assume that $\chi(\vec{q},\omega)$ is given]. Taking advantage now of the fact that perturbations are weak, one finds that the system has an exact, yet relatively compact, analytical solution (quoted here for brevity in the limit $s\to 0$). For incident LA waves (A_1 specified; $A_2=0$), the far-field ($z<-h$) LA amplitude in reflection is

$$A_3=-\Lambda_1\exp(-2ik_lh)+R_{ll}(A_1-\Lambda_1)-R_{tl}\Lambda_2, \quad (6)$$

where Λ_1 and Λ_2 are the LA and TA amplitudes, respectively, contributed by the electron layer,

$$\Lambda_1=[Bn_2(\vec{q})/2\rho c_l^2\cos\theta]\exp(ik_lh),$$
$$\Lambda_2=[Cn_2(\vec{q})/2\rho c_t^2\cos\varphi]\exp(ik_th). \quad (7)$$

Here $B=a\sin^2\theta+b\cos^2\theta$ and $C=(a-b)\sin\varphi\cos\varphi$; $n_2(\vec{q})$ is the response (to zeroth order in $\vec{u}$) and

c_l, c_t are the sound velocities for LA and TA waves, respectively. In (6) R_{ll} and R_{tl} are amplitude reflection coefficients—i.e., ratios of reflected amplitude to incident amplitude—in an obvious notation. The fractional change in reflected intensity for LA modes is

$$\Delta I(q)/I = 2\,\mathrm{Re}\{(A_3^0)^* \delta A_3\}/|A_3^0|^2, \tag{8}$$

where A_3^0 is the unperturbed amplitude and $\delta A_3 = A_3 - A_3^0$. Substituting (6) and (7) into (8) we obtain the final result (for limit $s \to 0$) to first order in $\chi(\vec{q}, \omega)$,

$$\begin{aligned}\frac{\Delta I(q)}{I} = &-\frac{1}{R_{ll}}\frac{B^2}{\rho c_l^2 \cos\theta \sin\theta}[q\,\mathrm{Im}\chi \cos 2k_l h - q\,\mathrm{Re}\chi\, \mathrm{sim}2k_l h] - \frac{2B^2}{\rho c_l^2 \cos\theta \sin\theta} q\,\mathrm{Im}\chi \\ &+\frac{R_{tl}}{R_{ll}}\frac{BC}{\rho c_t^2 \cos\varphi \sin\theta}[q\,\mathrm{Im}\chi \cos(k_t - k_l)h + q\,\mathrm{Re}\chi \sin(k_t - k_l)h] \\ &-\frac{R_{lt}}{R_{ll}}\frac{BC}{\rho c_l^2 \cos\theta \sin\varphi}[q\,\mathrm{Im}\chi \cos(k_t - k_l)h + q\,\mathrm{Re}\chi \sin(k_t - k_l)h] \\ &+\text{four terms of higher order in } R. \end{aligned} \tag{9}$$

A plot of $\Delta I(q)/I$ (all eight terms and finite s) is shown in Fig. 2(a) for $\chi(\vec{q}, \omega)$ calculated for a noninteracting electron gas[7] at 0 K [Fig. 2(b)].

Despite its complicated appearance (9) has a straightforward interpretation: First, the interference we anticipated between waves radiated from the 2D electron gas and waves reflected from the Si-SiO_2 interface is evidenced by its oscillatory nature depending upon arguments which represent the phase difference, $\Delta\varphi = \varphi_{\mathrm{rad}} - \varphi_{\mathrm{refl}}$, between the respective waves. Clearly "absorption" is a misnomer; $\Delta I(q)/I$ could be either positive or negative [see Fig. 2(a)] according to the magnitude of $2h$ relative to an inverse (perpendicular) phonon wave vector k^{-1}. Since $2h \lesssim 2\pi k^{-1}$ for all relevant k's, the oscillation is limited to one cycle or less. If $2h$ were large, $\Delta I(q)/I$ would appear as rapid oscillations modulated by $q\,\mathrm{Im}\chi$ and $q\,\mathrm{Re}\chi$. Second, in contradiction to the static case where only $\mathrm{Im}\chi$ appears to represent dissipation, it is seen that *both* $\mathrm{Re}\chi$ and $\mathrm{Im}\chi$ appear in (9) in a mix depending upon $\Delta\varphi$. The inclusion of $\mathrm{Re}\chi$, which is larger than $\mathrm{Im}\chi$ [see Fig. 2(b)], gives an enhancement to $\Delta I(q)/I$ as well as contributions beyond $2k_F$. Third, there is a substantial enhancement, by $|R_{ll}^{-1}| \sim 5$, of the leading term (successive terms are grouped in ascending order of number of reflections, i.e., in descending magnitude). The origin of each term is identified in Fig. 3. The first, zeroth order in reflection number, is from direct radiation (wave 1) excited by the incident LA waves. The next three terms are first order in reflection number: The second is bona-fide absorption[8] suffered by incident and reflected primary waves (hence, the prefactor "2") in passage through the electron layer (absorption is represented by interference of out-of-phase radiated waves 2 and 2′ with the primary waves). The third and fourth terms involve TA waves mode converted at the

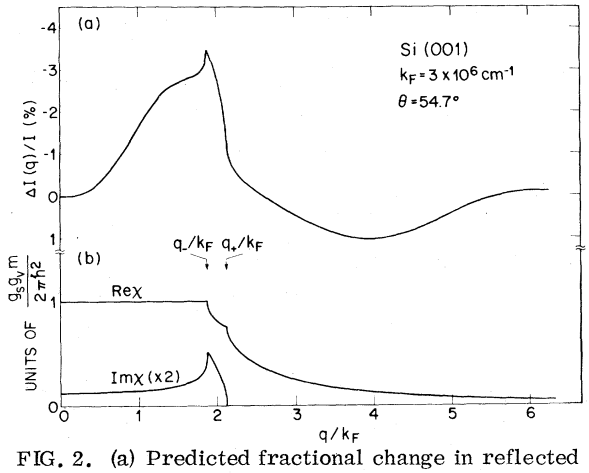

FIG. 2. (a) Predicted fractional change in reflected intensity $\Delta I(q)/I$ for LA phonons of wave vector q in the plane. Parameters used are $\rho = 2.33\ \mathrm{g/cm^3}$, $c_l = 9.4 \times 10^5$ cm/sec, $c_t = 5.11 \times 10^5$ cm/sec, $s = 2a_0$, $h = 4.5a_0$, $R_{ll} = -0.2$, $R_{lt} = 0.15$, and $R_{tl} = -0.1$. Designated q's are $q_\pm = 2k_F \pm 2mc/\hbar$, where $c = c_l/\sin\theta = c_t/\sin\varphi$. (b) $\chi(\vec{q}, \omega)$ for noninteracting electrons of valley degeneracy $g_v = 2$, spin degeneracy $g_s = 2$, and effective mass $m = 0.19m_0$.

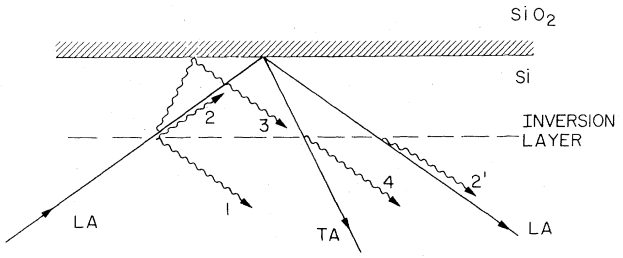

FIG. 3. Identification of individual contributions to $\Delta I(q)/I$ [cf. Eq. (9) in text].

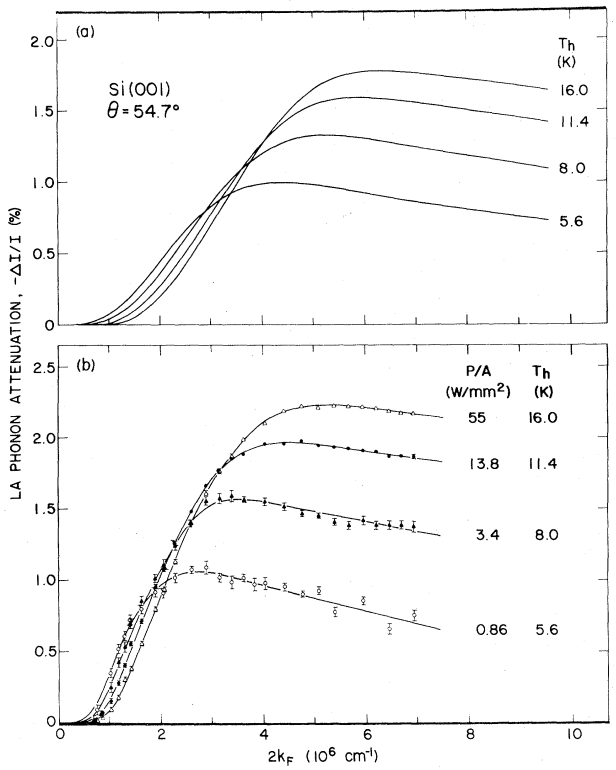

FIG. 4. (a) Profiles of $\Delta I/I$ for LA phonons in reflection obtained by the convolution of $\Delta I(q)/I$ [e.g., Fig. 2(a)] with phonon distributions characterized by a heater temperature T_h. (b) Experimental data from Ref. 2.

interface (waves 3 and 4, respectively). The four omitted terms are second order in reflection number. All told, "true" absorption accounts for only ~10% of $\Delta I(q)/I$, the remainder being a "reactive" response.

The experiments were performed with a *spectrum* of phonons, so that (9) must be convolved with an appropriate distribution function.[2] The resulting profiles, plotted versus $2k_F$ in Fig. 4(a), show good agreement with the data in the main areas where accord was formerly lacking: Most importantly, the calculated attenuation is now of the right magnitude; even the range (1% to 2%) is found to match roughly. Also there is now much better agreement in the shapes, in particular the crossing of curves at small $2k_F$ and the elevated tails of all profiles at large $2k_F$. [The latter are largely due to the added strength at small q's as seen in Fig. 2(a).] The only serious discrepancy is that the curves are shifted to higher $2k_F$ values for reasons that are obscure at this time. Experiments with monochromatic phonons may be necessary to identify whether the problem is in the theory or in our assumptions regarding the phonon distributions.

The parameters used in the computations are taken as "given" (see caption of Fig. 2) and not adjusted. The only exceptions are the parameters s and h which because of their somewhat artificial nature are not physically definable in a precise way. We find that the profiles show rather little dependence on s (if the midposition of the layer at $z = -h + s/2$ is held fixed). On the other hand, since interference is the major effect, the dependence on h is strong. If we write $h - s/2 = \text{const} \times a_0$ where a_0 is the variational thickness parameter given by Eq. (3.30) of Ref. 3, the best fit occurs when const = 3.5, rather close to the variational first moment of z for the charge density, $z_0 = 3a_0$.

In conclusion, the reconcilement of theory with the observed magnitude of $\Delta I/I$ demonstrates that the strength of the electron-phonon interaction in 2D does not differ greatly from that in 3D. Furthermore, with monochromatic phonons it will be feasible under appropriate conditions to do direct spectroscopy on the linear response function $\chi(\vec{q}, \omega)$ over a substantial region of $\vec{q}$ and ω, something that is otherwise fairly inaccessible. This might be particularly interesting in the presence of a quantizing magnetic field.

One of the authors (B.I.H.) is grateful for the hospitality of the Institute of Theoretical Physics, Santa Barbara. This work was supported in part by the National Science Foundation through Grant No. DMR 82-07431.

[1]J. C. Hensel, R. C. Dynes, and D. C. Tsui, Surf. Sci. 113, 249 (1982), and J. Phys. (Paris), Colloq. 42, C6-308 (1981).

[2]J. C. Hensel, R. C. Dynes, and D. C. Tsui, Phys. Rev. B 28, 1124 (1983).

[3]T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. 54, 437 (1982).

[4]Disordered SiO_2 is a strongly absorbing medium for ballistic phonons [see W. Dietsche and H. Kinder, Phys. Rev. Lett. 43, 1413 (1979)], so that the ~90% fraction transmitted across the interface is essentially dissipated in the thick (8000-Å) oxide layer.

[5]G. L. Bir and G. E. Picus, *Symmetry and Strain-Induced Effects in Semiconductors* (Wiley, New York, 1974), p. 306.

[6]In Herring's notation, $a = \Xi_d$ and $b = \Xi_d + \Xi_u$. For Si $\Xi_d = -6.0$ eV and $\Xi_u = +9.0$ eV [see, e.g., K. Murase, K. Enjouji, and E. Otsuka, J. Phys. Soc. Jpn. 29, 1248 (1970)].

[7]F. Stern, Phys. Rev. Lett. 18, 546 (1967).

[8]This one term comprises the entire result of the earlier theory (Refs. 1 and 2) except for a form factor. Note that $\pi G(q) = q\,\mathrm{Im}\chi$ where $G(q)$ is the absorption function defined in Ref. 2.