CARRIER LOCALISATION IN INVERSION LAYERS
AND IMPURITY BANDS

ANNUAL TECHNICAL REPORT

by

M. PEPPER

November 1980

EUROPEAN RESEARCH OFFICE

United States Army

London England

GRANT NUMBER DA-ERO - 78-G-098

GR:

Cavendish Laboratory
Department of Physics
University of Cambridge
Cambridge
CB3 OHE
U.K.

Approved for Public Release; distribution unlimited
**Title:** Carrier Localisation in Inversion Layers and Impurity Bands

**Authors:** M. Pepper

**Abstract:**

This report contains results on the ballistic injection of electrons between A1 and n+ Si and A1 and the Si inversion layer. This was the first experiment on ballistic injection between a semiconductor and a metal, and yielded results of considerably greater clarity than those found from the all metal work. Basically the phenomena responsible for intervalley scattering can be clearly observed and the technique offers a means of...
20. Contd.

identifying the various subbands in the Si inversion layer. Plasmons could be observed and in future it is planned to extend this technique to III-V semiconductors.

The other area of research described in this report is a collaborative experiment with Dr. K. von Klitzing and Dr. G. Dorda. It is shown that the Hall resistance of a MOSFET, when the conduction band is quantized by a strong magnetic field, is related simply to the fine structure constant. The experiment is described in detail.
ABSTRACT

This report contains results on the ballistic injection of electrons between Al and n+ Si and Al and the Si inversion layer. This was the first experiment on ballistic injection between a semiconductor and a metal, and yielded results of considerably greater clarity than those found from the all metal work. Basically the phonons responsible for intervalley scattering can be clearly observed and the technique offers a means of identifying the various subbands in the Si inversion layer. Plasmons could be observed and in future it is planned to extend this technique to III-V semiconductors.

The other area of research described in this report is a collaborative experiment with Dr. K. von Klitzing and Dr. G. Dorda. It is shown that the Hall resistance of a MOSFET, when the conduction band is quantized by a strong magnetic field, is related simply to the fine structure constant. The experiment is described in detail.
KEYWORDS

Silicon

Silicon - Silicon Dioxide Interface

Phonons

Localization

Ballistic Injection

Low Temperatures
CONTENTS

1. Ballistic injection of electrons in metal-semiconductor junctions
   1a Phonon spectroscopy and impurity enhanced inelastic scattering in $n^+$ silicon
   1b Phonon spectroscopy of aluminium
   1c Phonon and subband spectroscopy of silicon inversion layers

2. A new method for high accuracy determination of the fine structure constant based on quantized Hall resistance
1. Ballistic injection of electrons in metal-semiconductor junctions

1a Phonon spectroscopy and impurity enhanced inelastic scattering in $n^+$ silicon

It has been known for some time that the electron-phonon coupling in metals can be investigated by the use of point contacts. If the contact area is sufficiently small, structure is found in the I-V characteristic. Work by Yansen\(^1\),\(^2\) and Jansen et al.\(^3\),\(^4\) showed that \(d^2I/dV^2\) proportional to the electron-phonon coupling constant, \(g(\omega)\), given by

\[
g(\omega) = a^2(\omega) F(\omega).
\]

Here \(a^2(\omega)\) is the square of the matrix element for electron-phonon coupling averaged over the Fermi surface, and \(F(\omega)\) is the phonon density of states. These authors have investigated \(g(\omega)\) for a number of metals.

The theory of electrical resistance arising from small contact areas was first considered by Maxwell\(^5\), who predicted a resistance \(R\) given by \(1/2\sigma b\), \(\sigma\) is the bulk conductivity and \(b\) is the contact radius. Sharvin\(^6\) considered the ballistic nature of current flow which occurs when the contact radius is smaller than the bulk mean free path, \(\ell\). He derived the contact resistance, \(R\), in the following manner. The current flowing through the junction, \(I\), is given by

\[
I = n e b^2 \delta v.
\]

\(n\) is the carrier concentration and \(V\) is the velocity acquired by a carrier on passing through the junction. If the Fermi energy, \(E_F\) \(>>\) \(eV\), \(v\) is given by \(eV/2m_F\), where \(V\) is the applied voltage and \(V_F\) is the Fermi velocity. Thus, \(R = 2m_F/n\sigma b^3 = 2\ell/\pi\sigma b^3\).

Wexler has given an interpolation formula,

\[
R = \frac{(K)}{2\sigma b} + \frac{4K}{3\pi\sigma b}.
\]

Here \((K)\) is a function of the Knudsen ratio, \(\ell/b\), being 1 for \(K = 0\) and \(\sim 0.7\) for high \(K\). The first term in equation 2 is a "Maxwell" term and the second a "ballistic" term, i.e. independent of the bulk mean free path. Theoretical work, for example van Gelder\(^7\), has shown that minima in \(d^2I/dV^2\) are due to electrons which are initially emitted but then scattered back into the contact region. Each minima corresponds to the onset of a particular scattering process. Van Gelder finds that

\[
\frac{d^2I}{dV^2} = \frac{5.3 e \beta \theta}{4V_F} g(eV).
\]

Here \(e\) and \(\beta\) have their usual meanings, \(\beta\) is a factor close to 1 for high Fermi energies and \(\theta\), which is near unity, is the transmission factor of the contact. It was found that the background in \(d^2I/dV^2\) arises from multiple phonon emission processes. All the previous work
on this topic utilized metal-metal junctions, and, in general, the principal peaks in the phonon density of states corresponded to minima in $d^2I/dV^2$. In this work the first results were obtained on ballistic emission between a metal point contact ($\mathcal{A}$) and a degenerate semiconductor (Si). As will be shown the various phonon scattering processes are observed more clearly than for emission in metal-metal junctions.

The structures used comprised a thin film of $\mathcal{A}$, of thickness about 1 $\mu$m, separated from an $n^+$ diffusion in Si by a 500 $\AA$ - 1000 $\AA$ thick thermally grown film of $\text{SiO}_2$. A filamentary contact, ("short"), between $\mathcal{A}$ and the $n^+$ Si was formed by the application of a high field. For values of applied voltage up to 100 mV, the 4.2 K contact resistance was between 5 k\(\Omega\) and 10 k\(\Omega\) decreasing slightly with increasing voltage.

An important difference between this system and the metal-metal system is that there is here a considerable disparity between the conductivity of the $\mathcal{A}$' and the conductivity of the $n^+$ Si of doping $1-3 \times 10^{20}$ cm$^{-3}$. This will result in the voltage drop occurring in the Si. Although the $\mathcal{A}$ filament is probably polycrystalline the resistance of the filament was estimated to be small compared to the observed resistance. From equation 2 the radii of the contacts used in this work were estimated to be in the range 20 $\AA$ - 50 $\AA$. It was possible to observe a small hole in the $\mathcal{A}$ on the surface of the oxide where the filament had formed. This was about 1 $\mu$m in diameter, indicating that the filament narrows rapidly, or that the radius of contact is only a fraction of the total radius.

It is to be noted that this type of experiment is only possible in the absence of tunnelling, i.e. a depletion region is not present at the contact. This implies the use of $n^+$ doping and junctions with very small barrier heights. Unfortunately, it does not seem possible to use equation 1 for the extraction of the coupling constant $g(\omega)$. This is because most results are obtained for values of $V$ where $eV^2/\epsilon_F$, $\epsilon_F$ being 200 meV for the $n^+$ Si used in this work. Even where $eV \gg \epsilon_F$, $\epsilon$ does not increase as $V^1$, but faster than $V$, indicating that, perhaps, space charge effects are important in the depleted Si around the junction region.

Figure 1 shows $d^2I/dV^2$ plotted against $V$ for injection into arsenic doped Si at 4.2 K, this is achieved by biasing the Si positive with respect to the $\mathcal{A}$. The major peaks are identified in the figure. Essentially, the strongest effects are associated with the T.A. phonon responsible for $g$ scattering, i.e. between valleys on the same (100) axis, the g T.O. and, or, intravalley phonons. The low energy g process is forbidden in zero order by group theory and occurs as a higher order process. It is to be stressed that the combination minima do not correspond to the simultaneous emission of phonons, as here we would expect the strength of the minima to decrease rapidly with increasing number of phonons emitted. Rather they are due to electrons which have emitted $n$ phonons, where $n = 0, 1, 2, ...$, over a period of time then emitting another and being scattered back into the junction.

Figure 2 shows the corresponding result for injection into phosphorus doped Si. The principal points are the reduction in strength of the lowest energy g process, the virtual disappearance of the major peak at 30.5 meV, and the appearance of new structure near 40 meV. Table 1 contains a summary of the results and compares them with those obtained from magneto-phonon resonance studies on high purity Si (Raves et al.).
From a qualitative point of view, it is reasonable that the heaviest impurity, As, enhances the lower energy modes. It is relevant that impurity resonances have been observed in infra-red absorption at 54.6 meV, (P), and 65.3 meV(As), far from the structure reported here. The theory of absorption has been discussed by Bäuer and Elliott, and their predicted optical resonances are also distant from the impurity aided scattering processes found in this work.

We propose the following explanation of the additional structure due to impurities. The new phonons which assist inter-valley transitions are those found at the K point, (0.75, 0.75, 0), on the phonon dispersion curve. The scattering is an f process, i.e., between (100) and (010) valleys, and conservation of momentum is maintained by the electron transferring momentum to the impurity, which eventually emits a shower of low energy phonons. In addition, the presence of charged impurities will enhance the electron-phonon coupling. The As peak at 30.6 meV is associated with the TA phonon at K possessing E₃ symmetry. The E₃ phonon is not observed, being forbidden by time reversal symmetry. The transfer of momentum to the As likewise enhances the role of the g T.A. phonon which is normally weak. Phosphorus enhances the higher energy phonons near 40 meV, we associate these with the I.A. phonon at K with E₁ symmetry and the LO phonon at K with E₃ symmetry. Both these phonons are near 40 meV for pure Si, and possibly the shift in energy arises from a change in the force constants as well as the difference in mass between P and Si.

These differences between P and As were general but the strength of the modes, and the shape of the background, could differ between specimens prepared in different ways, with different values of doping. Figure 3 shows a case of injection into Si:P where d²I/dV² becomes negative; this appeared due to both strong Phosphorus enhancement and strong f scattering near 50 meV. This junction also showed evidence of oscillations, possible due to quantum effects, (when the modulation voltage was low), previously observed in metal-metal contacts (Jensen et al.). It is not clear if these are due to interference effects between multiple junctions or diffraction effects in single junction injection.

The principle features of the phonon spectrum were apparent at 77K, as shown for Si:As in Figure 4. It is noteworthy that the 30.5 meV minimum is not as dominant at this temperature, possibly absorption processes arising from the enhanced density of low energy phonons increase the importance of conventional intervalley scattering.

Pronounced structure was observed at high energies and is reproduced in Figure 5. This structure consists of a double resonance, each resonance is sharp and the two are separated by 48 meV. The change in d²I/dV² is, approximately, a factor of 30 greater than the change due to optical phonon coupling. The resonance is unlike the minima observed for phonon interactions, and indicates, that strong coupling is producing two minima in the electronic density of states. The energy and carrier concentration strongly suggest that the structure is due to a pair of plasmons, although the uncertainty in the carrier concentration prevents identification as a surface or bulk plasmon.

The maximum change in resistance of the electron gas is too small to allow plasmon enhanced intervalley scattering. Thus, this cannot be the origin of the splitting. This may result from valley-valley
interactions introducing a small, extra, anisotropy of the mass, or possibly, the effects of strain on surface plasmons. This point is being investigated in detail.

REFERENCES


FIGURE CAPTIONS

1. \( \frac{d^2 I}{dv^2} \) for electron injection into Si:As at 4.2 K. The principal phonons are indicated. The g(T.O.) and intravalley (I.V.) phonons are too close in energy to be separated.
2. \( \frac{d^2 I}{dv^2} \) for electron injection into Si:P at 4.2 K.
3. \( \frac{d^2 I}{dv^2} \) for electron injection into Si:P at 4.2 K. The top curve showing possible quantum effects was obtained with a modulation voltage of 0.1 mV, the lower curve with 1 mV. The L and T A\( \delta \) phonons are clear, these are discussed in the following paper.
4. \( \frac{d^2 I}{dv^2} \) for injection into Si:As at 77 K. The A\( \delta \) L phonon is clear as is the combination of this with the Q phonon. These are discussed in the following paper.
5. Structure in \( \frac{d^2 I}{dv^2} \) attributed to two plasmons. The magnitude of the structure is about a factor of thirty greater than the strongest phonon produced minimum.
TABLE I
PRINCIPAL INELASTIC SCATTERING PROCESSES IN SI.

<table>
<thead>
<tr>
<th>Si:P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy of Pheson</td>
</tr>
<tr>
<td>Si:As</td>
</tr>
<tr>
<td>Energy of Pheson</td>
</tr>
</tbody>
</table>

High Purity Si, obtained from Magneto-Pheson Resonance by Boys et al. (1975) = 3.

<table>
<thead>
<tr>
<th>Energy</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9</td>
<td>g phonon, T.A. mode, which is forbidden by group theory. From the phonon dispersion curve the energy should be 11.7 meV, assuming the conduction and valence bands to be at (0.83, 0.5). Coupling is weak in high purity Si and Si:P but strong in Si:As.</td>
</tr>
<tr>
<td>11.3</td>
<td>12.25</td>
</tr>
<tr>
<td>11.5</td>
<td>12.25</td>
</tr>
<tr>
<td>17.3</td>
<td>19.3</td>
</tr>
<tr>
<td>18.8</td>
<td>20.3</td>
</tr>
<tr>
<td>20.5</td>
<td>This is small in Si:P, very strong in Si:As. The suggested origin, described in the text, is an acoustic phonon at the K point with ( \Gamma_3 ) symmetry.</td>
</tr>
<tr>
<td>23.3</td>
<td>Suggested origins are the ( \Gamma_1 ) acoustic and ( \Gamma_3 ) optical modes at K. These phonons are weaker in Si:As.</td>
</tr>
<tr>
<td>43.9</td>
<td></td>
</tr>
</tbody>
</table>
TABLE I contd.

<table>
<thead>
<tr>
<th>Si:P</th>
<th>Si:As</th>
<th>High Purity Si, obtained from Magneto-Phonon Resonance by Daves et al. 1975 ± 3</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy of Phonon = 40.3 MeV</td>
<td>Energy of Phonon = 48.5 MeV</td>
<td>46.4</td>
<td>f phonons, strong in both Si:As and Si:P, they are a mixture of optical and acoustic modes.</td>
</tr>
<tr>
<td>51.5</td>
<td>51.0</td>
<td>50.8</td>
<td></td>
</tr>
<tr>
<td>60.4</td>
<td>60.0</td>
<td>60</td>
<td>f optical phonon, predicted to be very close to 60 meV. It is stronger in Si:P than Si:As.</td>
</tr>
<tr>
<td>65.5</td>
<td>65.0</td>
<td>65.9</td>
<td>Either, or both, the g T.O. phonon at 65.1 meV and an intravalley T.O. phonon at a similar energy.</td>
</tr>
<tr>
<td>&gt;65.5</td>
<td></td>
<td></td>
<td>A number of strong minima were observed, particularly at 65 and 66 meV. These could be fitted to combinations of the strongest phonons in the system, f phonons at 48.5 and 51.0 meV and Phosphorus enhanced phonons at 60.3 and 62.5 meV.</td>
</tr>
<tr>
<td></td>
<td>72.7</td>
<td></td>
<td>A combination of the 60 meV phonon with a low energy g phonon.</td>
</tr>
<tr>
<td>Energy of Phonon (meV)</td>
<td>Silicon Doped by Hg</td>
<td>High Purity Si, obtained from Manipo-Parson Resonance by Buss et al. 1975</td>
<td>Comments</td>
</tr>
<tr>
<td>------------------------</td>
<td>---------------------</td>
<td>------------------------------------------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>22.4</td>
<td></td>
<td>A combination of the 11.3 meV g phonon and either the T0 or the intervalley optical phonons.</td>
<td></td>
</tr>
<tr>
<td>29.7</td>
<td></td>
<td>A combination of the 11.3 meV g phonon and either the T0 or intervalley optical phonons.</td>
<td></td>
</tr>
<tr>
<td>320</td>
<td></td>
<td>A broad minimum with considerable structure due to combinations of optical phonons and the 30.5 meV line.</td>
<td></td>
</tr>
<tr>
<td>320 meV</td>
<td></td>
<td>2 large minima separated by 8 meV, possibly conduction electron phonon resonance.</td>
<td></td>
</tr>
</tbody>
</table>
INJECTION FROM AL INTO Si:As

\[ \frac{d^2I}{dV^2} \]

ARBITRARY UNITS

\[ V(\text{mV}) \]

\[ g(\text{L.A.}) \]
\[ g(\text{I.V.}) + 2g(\text{T.A.}) \]
\[ g(\text{T.O.}) \]
\[ g(\text{T.O.}) + g(\text{T.A.}) \]

\[ \text{As}, (\text{T.A. at K}) \]
INJECTION INTO Si:P, 4.2 K

\[ \frac{d^2i}{dV^2} \]

ARBITRARY UNITS

\[ g(L.A.) \]
\[ g(T.A.) \]
\[ g(T.O.) \]

PENHANCED

COMBINATIONS

V(mV)
In the previous section it was shown that electron injection into Si revealed the phonons responsible for scattering. When the $A_f$ is positive with respect to the Si, the $d^2I/dV^2$ against $V$ plot shows minima corresponding to $A_f$ phonons, Figure 1. The clarity of this figure, and the generally clear observation of $A_f$ phonons, contrasts with the reported failure of $A_f$-$A_f$ junctions.

Inspection of the $E$-$k$ diagram of $A_f$, (for example Harrison, shows that there are many possible types of phonon scattering. The most important processes are, a) between $X$ points; b) between $W$ points, $(W = 0, 0, 1)$; c) between $U$ points, $(U = 0, 1, 1)$; d) between $X$ and $W$; e) between $X$ and $U$; f) between $W$ and $U$, and g) scattering between the two bands near $U$. The principal phonons involved in these processes have coordinates, a) $(100)$ - an Umklapp process; b) $(0, 0, 1)$ for both normal and Umklapp processes; c) $(0, 0, 1)$; d) $(0, 0, 1)$, for normal and Umklapp processes, c) $(1, 1, 0)$ and $(1, 1, 0)$ for Umklapp processes; f) $(1, 0, 0)$, Umklapp process; g) this is a phonon close to the $(100)$ direction, and, assuming only the component in this direction, has coordinates $\gamma$ $(17, 0, 0)$. From the $A_f$ phonon dispersion curves, the phonons responsible are a) longitudinal, ($L$), and transverse, ($T$), phonons at the $X$ point having energies of 37.5 meV and 22.5 meV; b) the ($L$) phonon with energy 35 meV, this phonon gives rise to the principal peak in the density of states. There are also two transverse phonons with the same momentum, $T_1$ at 19 meV and $T_2$ at 25 meV; c) the ($L$) phonon at the $K$ point with energy 29 meV and two ($T$) phonons, $T_1$, with a high density of states at 21 meV, and $T_2$ at 35 meV.

Thus scattering is accomplished principally by 35 meV $L$ phonons and 21 - 22 meV $T$ phonons. A range of phonons can cause processes d) - f), which will make identification difficult. On the other hand, process g) will be caused by low energy phonons and should be well defined. Inspection of the phonon dispersion curves shows that the possible phonon energies for process $g$ are, approximately, an $L$ mode at 11 meV and a $T$ mode at 6 meV. In Figure 1 these are identified as the minima $P$ and $Q$ at 6 meV and 12 meV respectively. Minimum $R$ at 26 meV is identified as the $T_2$ phonon at $(1, 0, 0)$. A minimum found at 17 meV is attributed to the $T$ phonon having coordinates (111), i.e. at the $\Gamma$ point.

The major features of Figure 1 are clearly associated with the 35 meV $L$ and 22 meV $T$ phonons. In view of the many subsidiary minima, and the large number of possible phonon combinations, a detailed analysis of the figure was not performed. As seen in Figure 1, phonon combinations are clearly observed. It is not clear why combinations of $L$ phonons fade out abruptly, or why the $T$ combinations increase in strength above 160 meV. The degree of electron heating appears too small to sample additional features of the band structure. The combinations were observable for most of the specimens used in this work, except for that used for Figure 3 of the preceding section, where structure was not found past the principle $L$ minimum.

The principle features of the $A_f$ observations were clear at 77 K, for example, the $A_f$ side of Figure 2 in the previous paper clearly shows the 35 meV $L$ phonon, and, surprisingly, its combinations with the 12 meV $Q$ phonon thought responsible for scattering near the $U$ point.
It is noteworthy that this experiment has revealed a clearer $\text{Al}$ phonon spectrum than tunnelling experiments. Possibly the $\text{Al}$ filament is ordered by local heating during the formation process, or the electrons pass through the filament into the evaporated $\text{Al}$ before being scattered back.

REFERENCES


FIGURE CAPTIONS

1. The plot of $d^2I/dV^2$ against $V$ for injection into $\text{Al}$. The increase in low energy structure caused by decreasing the modulation voltage from 2mV (---) to 0.6 mV (---) is clear. The principal phonons and combinations are indicated and are discussed in the text. Sometimes those indicated were minima only at low modulation voltages, and shoulders at the higher modulation voltage. A wide range of combinations was found for a low modulation voltage. This was not analysed in detail.

Phonon and subband spectroscopy of silicon inversion layers

For this work MOSFET's were fabricated on the (100) surface of 10 $\Omega$ cm p type Si, ($N_A = 2.10^{15}$ cm$^{-3}$). A filamentary contact was formed between the $\text{Al}$ gate and the Si through 800 A of thermally grown $\text{SiO}_2$. In order to sustain an inversion layer without applying a voltage to the gate, the $\text{SiO}_2$ was contaminated with Na$^+$ ions before evaporation of the $\text{Al}$. The Na$^+$ was drifted to the Si-$\text{SiO}_2$ interface, and the device was cooled to 77 K to freeze the Na$^+$ into position prior to formation of the filamentary contact. Unfortunately, at this temperature the resistance between the inversion layer and the substrate was too high, and the contact was always formed between the gate and the source or drain. The contact was formed satisfactorily at 300 K, but at this temperature some drift of the Na$^+$ occurred and the carrier concentration could not be defined more accurately than $2.5 \pm 0.5 \times 10^{12}$ cm$^{-2}$. Although the Si is p type, $E_g$ at the Si surface is in the conduction band tail, and so a barrier does not exist between the $\text{Al}$ and the inversion layer.

It was found that at 4.2 K the resistance of the inversion layer in series with the contact was too high for meaningful measurements to be taken. Consequently the experiment was performed at 77 K where
the resistance was considerably lower. A number of minima were found on the $d^2I/dV^2$ versus gate voltage plot for injection into the inversion layer; the principal L minimum was clearly observed for injection into the Al.

Before considering the phonon spectrum in the inversion layer, we first briefly discuss phonon scattering in this system. The inversion layer differs from bulk Si in that the degeneracy of the valleys is lifted by the strong surface field. For the (100) surface, the ground state, $E_0$, is formed from the two valleys with heavy mass perpendicular to the interface; there is also a series of excited states associated with these two valleys, $E_1$, $E_2$, $E_3$, ... In addition, another ladder of energy levels is associated with the other four valleys which have light mass normal to the interface, $E_0'$, $E_1'$, $E_2'$, ... Ando has found that $E_0'$ lies close to $E_1'$.

Intervalley scattering in the twofold degenerate levels is by $g$ phonons. As there will be a sharp threshold for this process, it is observable despite the general background of intervalley scattering imposed by the absence of momentum conservation at the interface. For intrasubband transitions, the mobility is not affected by direct scattering along the major (100) axis. However, in systems with small $k_F$, phonons very close to the $q$ energy will scatter effectively. As the energy of the injected electrons increases we would expect to observe $g$ scattering in $E_0'$, followed by $g$ scattering in $E_1'$, and both $g$ and $f$ scattering in $E_0''$ as well as $f$ scattering between $E_0'$ and both $E_0$ and $E_1$. Scattering between subbands formed from the same set of valleys is forbidden in the zeroth order. Scattering across the Fermi surface by long wavelength phonons will not be observed.

The minima in $d^2I/dV^2$ as a function of $V$ were found and identified as particular phonons. However, two weak minima were found which could not be ascribed to phonons. If these were identified as the bottom of new subbands, a whole new phonon series followed in a rational way. These minima were at 29 meV and 55 meV. The energies are above the Si Fermi level, which cannot be ascertained directly due to a lack of precise knowledge of carrier concentration. Ignoring the existence of band tailing, and assuming a carrier concentration of $2.5 \times 0.5 \times 10^{12}$, we find that $E_F$ is roughly 12 meV. Thus the subbands are 241 meV and 57 meV above the bottom of $E_0$. These energies appear too great for $|E_1 - E_0|$ and $|E_2 - E_0|$ for the value of substrate doping. In fact the phonon spectrum suggested that the subbands are $E_0'$ and $E_1'$. This was not surprising as $E_0$ only gave one reasonable minimum - the 65 meV $g$ phonon. The details of $E_1$ and $E_2$ appear lost in the complex spectra of $E_0'$ and $E_1'$. The cause of the observation of weak minima corresponding to the onset of subbands must be in the greater density of states enhancing the probability of backscattering by phonon absorption. A very weak minimum was identified at ~7 meV, possibly this is due to surface (Rayleigh) phonons which may be important in determining the low temperature mobility.

Turning now to the detailed results shown in Table 1. It was barely possible to identify the low energy $g$ phonons although the 65 meV $g$ phonon was clear. A minimum was also observed at ~60 meV corresponding to an $f$ phonon. This could not arise from scattering within the $E_0$ band, but from electrons injected into $E_0'$ or $E_1'$ being scattered into $E_0$. 
### Table I

**Principai Phonons Participating in Inversion Layer Scattering, 77K**

<table>
<thead>
<tr>
<th>Subband at $E_0$</th>
<th>Energy of centre of minimum, meV ± 1.0 meV</th>
<th>Energy above bottom of subband, meV ± 1.0 meV</th>
<th>Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Ground subband, $E_0$</td>
<td>7</td>
<td>7</td>
<td>Possible surface mode</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>13</td>
<td>Extremely weak g phonons</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>18</td>
<td>f phonon scattering electrons into this subband from another subband</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>60</td>
<td>g or possibly intravalley phonon</td>
</tr>
<tr>
<td></td>
<td>63</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>b) Subband at 20 meV $E_0$</td>
<td>38</td>
<td>9</td>
<td>g phonons</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>11</td>
<td>g phonon</td>
</tr>
<tr>
<td></td>
<td>47</td>
<td>18</td>
<td>f phonon</td>
</tr>
<tr>
<td></td>
<td>52</td>
<td>24</td>
<td>f phonon, merged with an f phonon of the following series</td>
</tr>
<tr>
<td></td>
<td>78</td>
<td>49</td>
<td>f phonon</td>
</tr>
<tr>
<td></td>
<td>90</td>
<td>61</td>
<td>g or possibly intravalley phonon</td>
</tr>
<tr>
<td></td>
<td>96</td>
<td>67</td>
<td></td>
</tr>
</tbody>
</table>
TABLE I
contd.

<table>
<thead>
<tr>
<th>Subband at. 55 meV, E₁</th>
<th>Energy of Centre of minimum, meV</th>
<th>Energy above bottom of subband, meV</th>
<th>Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>c)</td>
<td>66</td>
<td>11</td>
<td>Very broad g phonon</td>
</tr>
<tr>
<td></td>
<td>77</td>
<td>22</td>
<td>g</td>
</tr>
<tr>
<td></td>
<td>78</td>
<td>23</td>
<td>f, possibly merged with an f phonon of preceding series</td>
</tr>
<tr>
<td></td>
<td>105</td>
<td>50</td>
<td>Appeared to be two partially resolved f phonons</td>
</tr>
<tr>
<td></td>
<td>108</td>
<td>53</td>
<td>f phonon</td>
</tr>
<tr>
<td></td>
<td>115</td>
<td>60</td>
<td>g or possibly intravalley phonon</td>
</tr>
<tr>
<td></td>
<td>122</td>
<td>67</td>
<td></td>
</tr>
</tbody>
</table>
Examination of the phonon series for the subband starting at 29 meV shows generally clear evidence of f scattering, supporting identification as E'₁. The electrons participating in this process may be scattered within E₀', or into this subband from E₀.

The subband commencing at 55 meV displays a comprehensive phonon spectrum, but it cannot be directly ascertained if this is of the E or E' series. The existence of f scattering is not conclusive as, in the absence of compression experiments, it cannot be determined if this is inter or intrasubband scattering. However, the strength of the g minima is more akin to E₀ than E₀', and so we tentatively identify this subband as E'

Further minima were observed at 128, 133 and 140 meV, and considerable structure was found at energies greater than 120 meV. Due to the complexity of identification, further analysis was not performed. These minima may result from higher subbands, phonon combinations or possibly plasmons.

In conclusion, this appears to be a promising technique for the investigation of both phonon scattering, which can be clarified by the application of compression, plasmon effects and subband spectroscopy. In this latter context, it is to be noted that it is possible to vary the subband splitting by the application of substrate bias.

REFERENCES

1. Ando, T., 1978 Surface Science, 73, 1
2. Ferry, D.K., 1975 Surface Science, 57, 218

A new method for high accuracy determination of the fine structure constant based on quantized Hall resistance*

In this section we report a new, potentially high-accuracy method for determining the fine-structure constant, a. The new approach is based on the fact that the degenerate electron gas in the inversion layer of a MOSFET (metal-oxide-semi-conductor field-effect transistor) is fully quantized when the transistor is operated at helium temperatures and in a strong magnetic field of order 15 T. The electric field perpendicular to the surface (gate field) produces subbands for the motion normal to the semi-conductor-oxide interface, and the magnetic field produces Landau quantization of motion parallel to the interface. The density of states D(E) consists of broadened δ functions; minimal overlap is achieved if the magnetic field is sufficiently high. The number of states, Nₗ, within each Landau level is given by

\[ Nₗ = \frac{e^2}{\hbar h} \]  (1)

*Experiments performed by G. von Klitzing of Stanford University, West Germany, as part of a collaborative project.
where we exclude the spin and valley degeneracies. If the density of
the states at the Fermi energy, \( N(E_F) \), is zero, an inversion layer
carrier cannot be scattered, and the centre of the cyclotron orbit
drifts in the direction perpendicular to the electric and magnetic
field. If \( N(E_F) \) is finite but small, an arbitrarily small rate of
scattering cannot occur and localization produced by the long lifetime
is the same as a zero scattering rate, i.e., the absence of
current-carrying states occurs. Thus, when the Fermi level is between
Landau levels the device current is thermally activated and the mini-
ma in \( \sigma_{xx}^{\text{min}} \), can be less than \( 10^{-7} \sigma_{xx}^{\text{max}} \). Increasing the magnetic
field and decreasing the temperature, further decreases \( \sigma_{xx}^{\text{min}} \). The
Hall conductivity \( \sigma_{xy} \), which is usually a complicated function of the
scattering process, becomes very simple in the absence of scattering
and is given by

\[
\sigma_{xy} = -\frac{Ne}{B}
\]

(2)

where \( N \) is the carrier concentration.

The correction term to the above relation, \( \Delta \sigma_{xy} \), is of the order
of \( \sigma_{xx}/\omega \tau \), where \( \omega \) is the cyclotron frequency and \( \tau \) is the relaxation
time of the conduction electrons; \( \omega \tau >> 1 \) in strong magnetic fields.
When the Fermi energy is between Landau levels, and \( \sigma_{xx}^{\text{max}} \sim 10^{-7} \sigma_{xx}^{\text{min}} \),
the correction \( \Delta \sigma_{xy}/\sigma_{xy} < 10^{-8} \). Subject to any error imposed by \( \Delta \sigma_{xy} \),
when a Landau level is fully occupied and \( N = N_L \) (i = 1, 2, 3 ...),
\( \sigma_{xy} \) is immediately given from equations (1) and (2):

\[
\sigma_{xy} = \frac{e^2 i}{h}.
\]

(3)

The Hall resistivity \( \rho_{xy} = -\sigma_{xy}/(\sigma_{xx}^2 + \sigma_{xy}^2) = -\sigma_{xy}^{-1} \) is defined by
\( R_H/j(E_H = \text{Hall field, } j = \text{current density}) \) and can be rewritten \( R_H/I \),
where \( R_H \) is the Hall resistance, \( U_H \) the Hall voltage and \( I \) the current.
Thus, \( R_H = h/e^2 i \), which may finally be written as

\[
R_H = a^{-1} \mu_0 c/2i,
\]

(4)

where \( \mu_0 \) is the permeability of vacuum and exactly equal to \( 4\pi \times 10^{-7} \)
H m\(^{-1} \), \( c \) is the speed of light in vacuum and equal to 299 792 458 m s\(^{-1} \)
with a current uncertainty of 0.004 ppm and \( a = 137 \) is the fine-
structure constant. It is clear from equation (4) that a high-accuracy
measurement of the Hall resistance in SI units to a few parts in \( 10^8 \)
by means of the so-called calculable cross capacitor by Thompson and
Lampard\(^5\), the question of absolute units versus as-maintained units is
much less of a problem than in the determination of \( e/h \) from the ac
Josephson effect. Furthermore, the magnitude of \( R_H \) falls within a
relatively convenient range: \( R_H \% (25 813 \Omega)/i \), with \( i \) typically
between 2 and 8. Finally, we note that if \( a \) is assumed to be known
from some other experiment (for example, from \( 2e/h \) and the proton
gyromagnetic ratio \( \gamma_p \) ), equation (4) may be used to derive a known
standard resistance.

Two well-known corrections in the low-field Hall effect become
unimportant. The first is the correction due to the shunting of the
Hall voltage by the source and drain contacts\(^7\). This is important at
low fields for samples with length-to-width ratio, \( L/W \), less than 1,
but becomes negligible when the Hall angle is \( 90^\circ \), i.e., \( \theta \approx 0 \).\(^8\)

The second correction which becomes unimportant is that due to an
inexact alignment of the Hall probes, i.e., they are not exactly oriented. This is irrelevant, as the voltage drop along the sample vanishes when \( \varepsilon_{xx} = 0 \).

The experiments were carried out on \( \text{SiO}_2 \) devices with a range of oxide thicknesses \( (d_{\text{ox}} = 100 \text{ nm} - 400 \text{ nm}) \), and length-to-width ratios ranging from \( L/W = 25 \) to \( L/W = 0.65 \). All the transistors were fabricated on the (100) surface orientation and, typically, the \( p \)-type substrate had room-temperature resistivity of \( 10 \ \Omega \). The resistivity of helium temperature was higher than \( 10^4 \ \Omega \ cm \), and no current flow between source and drain around the channel could be measured. The long devices \( (L/W > 5) \) had potential probes in addition to the Hall probes.

The measured voltage \( U_{pp} \) is proportional to the resistivity component \( \rho_{xx} = \varepsilon_{xx} \varepsilon_{xy}^2 + \varepsilon_{xy} \varepsilon_{xx}^2 \). At gate voltages where the \( \varepsilon_{xy} \) is in the energy gap between Landau levels, minima in both \( \rho_{xx} \) and \( \rho_{xy} \) are observed. Such minima are clearly visible, and are identified, in Figure 1; the minima due to the lifting of the spin and the (two-fold) valley degeneracy are also apparent. The Hall voltage clearly levels off at those values of carrier concentration where \( \varepsilon_{xy} \) and \( \varepsilon_{xx} \) are zero. The values of \( U_{H} \) obtained in the regions are in good agreement with the predicted values, equation (4), if the error due to the \( 1-\omega \) input impedance of the X-Y recorder is taken into account. It was found that the value of \( U_{H} \) in the "steps" was, for constant current, independent of sample geometry and direction of magnetic field, provided that \( \varepsilon_{xx} \) was zero.

An area of possible criticism of the theoretical basis of this experiment is the role of carriers which are localized outside the main Landau level. Here we do not specify the localization mechanism, but the presence of localized carriers will invalidate both the relation \( N = N_{H} \) and equation (4). However, the experimental results strongly suggest that such carriers do not invalidate equation (4). At present there is both theoretical and experimental investigation of this type of localization. Ando has suggested that the electrons in impurity bands, arising from short range scatterers, do not contribute to the Hall current, whereas the electrons in the Landau level give rise to the same Hall current as that obtained when all the electrons are in the level and can move freely. Clearly this process must be carefully examined as an accompaniment to highly accurate measurements of Hall resistance.

For high-precision measurements we used a normal resistance \( R_{0} \) in series with the device. The voltage drop, \( U_{0} \), across \( R_{0} \), and the voltages \( U_{pp} \) and \( U_{pp'} \) across and along the device was measured with a high impedance voltmeter \( (R = 2 \times 10^{12} \) \( \Omega \)). The resistance \( R_{0} \) was calibrated by the Physikalisch Technische Bundesanstalt, Braunschweig, and had a value of \( R_{0} = 9999.69 \) \( \Omega \) at a temperature of \( 293 \text{K} \). A typical result of the measured Hall resistance \( R_{H} = U_{pp}/(I_{pp} - I_{pp'})/V_{pp} \) and the resistance \( R_{pp} = U_{pp}/V_{pp} \) between the potential probes of the device is shown in Figure 2; \( V_{pp} = 23.6 \text{ V} \) corresponds to the minima at \( V_{pp} = 0.7 \) in Figure 1, because the thickness of the gate oxide of these transistors differ by a factor of 3.6. The experimental arrangements was not sensitive enough to measure a value of \( R_{pp} \) of less than \( 10^8 \) \( \Omega \). The value found in the gate-voltage region \( 25 \text{ V} < V_{pp} < 72 \text{ V} \). The Hall resistance in this gate-voltage region had a value of \( 6.3 \times 10^8 \) \( \Omega \). This inaccuracy of
was due to the limited sensitivity of the voltmeter. It would
like to mention that most of the samples, especially devices with a
small length-to-width ratio, showed a minimum in the Hall voltage as
a function of \( V_g \) at gate voltages close to the left side of the plateau.
In Figure 2, this minimum is relatively shallow and has a value of 6452.87 \( \Omega \) at \( V_g = 23.30 \).

In order to demonstrate the insensitivity of the Hall resistance
on the geometry of the device, measurements on two samples with a
length-to-width ratio of \( L/W = 0.65 \) and \( L/W = 25 \), respectively, are
plotted in Figure 3. The gate-voltage scale is given in arbitrary
units, and is different for the two samples because the thickness of
the gate oxides are different. A gate voltage \( V_g = 1.60 \) corresponds,
approximately, to a surface carrier concentration where the first
degenerate Landau level, \( n = 0 \), is completely filled. Within
the experimental accuracy of 0.1 \( \Omega \), the same value for the plateau in
the Hall resistance is measured. The value for \( h/4e^2 = 6452.20 \pm 0.05 \Omega \)
expected value for the fine-structure constant is plotted in this figure, too. The decrease of the Hall resistance with
decreasing gate voltage for the sample with \( L/W = 0.65 \) originates
mainly from the shorting of the Hall voltage at the contacts. This
effect is most pronounced when the Hall angle becomes smaller than
90°. In the limit of small Hall angles, the Hall voltage is reduced
by a factor of 2 for the sample with \( L/W = 0.65 \).

The mean value of the Hall resistance for all samples investigated
was 6453.22\( \pm 0.10 \) \( \Omega \) for measurements in the energy gap between the
Landau levels \( n = 0 \) and \( n = 1 \) (corresponding to \( i = 4 \) in equation 4),
3226.62\( \pm 0.10 \) \( \Omega \) for measurements in the energy gap between Landau levels
\( n = 1 \) and \( n = 2 \) (\( i = 8 \)), and 12906.5\( \pm 1.0 \) \( \Omega \) for measurements in the energy gap between the spin split levels with \( n = 0 \) (\( i = 2 \)). These resistances
agree very well with the calculated values of \( h/4e^2 \) based on the recently
reported highly accurate value of \( \alpha^{-1} = 137.035963(15) \) (0.11 ppm).

Measurements with a voltmeter with higher resolution and a
calibrated standard resistor with a vanishing small temperature coefficient
at \( T = 25^\circ \) yield a value of \( h/4e^2 = 6453.17\pm 0.02 \) \( \Omega \) corresponding to
a fine-structure constant of \( \alpha^{-1} = 137.0553\pm 0.0004 \).

REFERENCES

1. For a review see for example: F. Stern, Crit. Rev. Solid State Sci.,
State Comm., 11, 345
2, 631
Instrum., 19, 366
FIGURE CAPTIONS

1. Recordings of the Hall voltage $U_H$, and the voltage drop between the potential probes, $U_{pp}$, as a function of the gate voltage $V_g$ at $T = 1.5$ K. The constant magnetic field (B) is 18 T and the source drain current, I, is 1 μA. The inset shows a top view of the device with a length of L = 100 μm, a width of W = 50 μm, and a distance between the potential probes of $L_{pp} = 130$ μm.

2. Hall resistance $R_H$, and device resistance, $R_{pp}$, between the potential probes as a function of the gate voltage $V_g$ in a region of gate voltage corresponding to a fully occupied, lowest ($n = 0$) Landau level. The plateau in $R_H$ has a value of 6453.3 ± 0.1 Ω. The geometry of the device was L = 400 μm, W = 50 μm, and $L_{pp} = 130$ μm; B = 13 T.

3. Hall resistance $R_H$ for two samples with different geometry in a gate-voltage region $V_g$ where the $n = 0$ Landau level is fully occupied. The recommended value $\hbar/4e^2$ is given as 6453.204 Ω.
FIGURE 1.
FIGURE 2.

B = 13.0 T
T = 1.8 K
$B = 13.9 \, T$
$T = 1.8 \, K$

- $L = 260 \, \mu m, W = 400 \, \mu m$
- $L = 1000 \, \mu m, W = 40 \, \mu m$

$\frac{h}{4e^2}$

Gate voltage $V_g$ / rel. units

Figure 3.