1. Publication NO *INPE-3483-PRE/726*	2. Version	3. Date *March, 1985*	5. Distribution ☐ Internal ☒ External
4. Origin Program *DTE/DCT* *MATER/FISUP*			☐ Restricted

6. Key words - selected by the author(s)
IMPURITY BANDS
INVERSION LAYERS
MOSFET

7. U.D.C.: *539.2*

8. Title *INPE-3483-PRE/726* *IMPURITY BANDS ASSOCIATED WITH MOS INVERSION LAYERS*	10. NO of pages: *08*
	11. Last page: *07*
	12. Revised by *Irajá N. Bandeira*
9. Authorship *I.C. da Cunha Lima* Responsible author	13. Authorized by *Nelson de Jesus Parada* *Director General*

14. Abstract/Notes

Aspects of the impurity band formation in Si/SiO_2 MOSFET's are reviewed from the theoretical point of view and some recent results are shown. Differences with bulk doped semiconductors are also pointed out. Possible theoretical efforts for the future are outlined.

15. Remarks

To be published in the Proceedings of the Second Brazilian School on Semiconductors.

IMPURITY BANDS ASSOCIATED WITH MOS INVERSION LAYERS

I.C. da Cunha Lima

Instituto de Pesquisas Espaciais, INPE
Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq
12200 São José dos Campos - SP - Brasil

Aspects of the impurity band formation in Si/SiO_2
MOSFET's are reviewed from the theoretical point of view and some
recent results are shown. Differences with bulk doped semiconductors
are also pointed out. Possible theoretical efforts for the future are
outlined.

It has been shown in other talks of this BSSP2 how a
dynamically two-dimensional electron gas can be created in both an
MOS structure and a semiconductor heterojunction.[1,2] In what concerns
transport properties, the system behaves differently for different
regimes of carrier density and temperature.[2] In the case of low carrier
concentration (small gate voltage) and low temperature, the departure
from a metallic behaviour in the 2-D gas of a MOSFET is taken as a
consequence of localization due to disorder.

It is well-known that Na^+ impurities always occur in the
oxide close to the Si/SiO_2 interface of a MOSFET.[3] This, besides
roughness of the interface, is the important source of disorder of the
electron gas. However, surface roughness becomes unimportant in the
case of small gate voltage. Impurity scattering mechanisms are
responsible for the band tailing of the electron subband. But, since an
attractive potential can always generate a bound state in a two-
-dimensional (2-D) system, in the presence of an inversion layer a
single Na^+ impurity inside the oxide will always have an electron bound
in the semiconductor. As the Na^+ concentration increases, the bound
state wavefunctions start to overlap giving rise to an impurity band,
i.e., a broadening of the bound state energy level.

In order to study experimentally the influence of those
Na^+ impurities in the electronic conduction of a MOSFET, samples are
produced in which the impurities are deposited in the oxides and
drifted by an electric field to lie close to the oxide-semiconductor
interface. In those samples the Na^+ concentrations are fairly well-
-known.[3] As a result of those experiments impurity band formation below
the first subband has been observed.

A simple model to describe the bound state of an electron
to a Na^+ impurity in a MOSFET was devised by Stern and Howard[4] (SH)
almost two decades ago. In a few words, that model comes out from the
use of the effective mass theorem and assumes a Thomas-Fermi screening
by the 2-D electron gas. The potential energy term of the single
electron Hamiltonian becomes

$$V(r) = -e^2/\overline{\kappa}\,(z_0^2 + r^2)^{1/2} + V_s(r) \; , \tag{1}$$

where $\overline{\kappa}$ is the effective dielectric constant, $\overline{\kappa} = (\kappa_{ox} + \kappa_{semic})/2$, z_0 is the distance from the Na^+ ion to the interface, which is parallel to the plane xy. The second term is due to the screening. The screening constant s, the inverse of the screening length in the SH theory, is independent of the Fermi level as soon as the sample is kept in the electrical quantum limit, i.e., only the first subband occupied. Then, the constant s depends only on the sample's characteristics and is treated as a variable parameter in the theory, the same as for the z_0.

In the limiting case of $s \to 0$ ($V_s \to 0$) and $z_0 \to 0$ the energy eigenvalues for the bound states are

$$\varepsilon_n = -(n - 1/2)^{-2}Ry^* \tag{2}$$

below the edge of the first subband. The binding energy of the ground state is $-4Ry^*$, where $1Ry^* = m^*e^4/2\hbar^{-2}$. In the case of the (100) Si/SiO_2 MOSFET, $1Ry^* = 42meV$. Eq. (2) states that the energy eigenvalues obey a 2-D hydrogen atom sequence. This is not exactly true, however, because in Eq. (1) the potential energy is not logarithmic as it should be in the strictly 2-D case.

For nonzero s and z_0, the single impurity Schrödinger equation must be solved numerically. Only the ground state is to be obtained. The excited states, if existing, are not important for the purpose of this matter.

As it was said before, when many impurities are present, overlapping of wavefunctions corresponding to different attracting centers can give rise to an impurity band formation. The situation here is quite different from the case of n-type doped semiconductors. In uncompensated doped semiconductors the impurity density equals the electron density (in the case of one ionized electron per impurity) and determines the Fermi level. In the metallic regime the screening of the

impurity potential may become strong enough in such a way that no bound state is allowed to exist. The impurities then play the role of structureless scattering centers for the electrons and the disorder is responsible for the band tailing of the conduction band. If the impurity concentration decreases, an impurity band separates from the conduction band.[5]

Neglecting electron-electron interaction the following Hamiltonian can be used to describe the many-impurity case:

$$H = - \frac{\hbar^2 \nabla^2}{2m^*} + \sum_{\{\ell\}} V(\vec{r} - \vec{R}_\ell) \, , \tag{3}$$

where $V(\vec{r} - \vec{R}_\ell)$, the same as in Eq. (1), is the potential energy of the electron due to an attractive center located at the point $\vec{R}_\ell$ in the plane xy. Next, the eigenstates of H are expanded in terms of the ground states of h_ℓ's,

$$h_\ell = - \frac{\hbar^2 \nabla^2}{2m^*} + V(\vec{r} - \vec{R}_\ell) \, , \tag{4}$$

$$h_\ell |\ell\rangle = \varepsilon_\ell^{(0)} |\ell\rangle \, . \tag{5}$$

The impurities are assumed to be all at the same distance from the interface, therefore $\varepsilon_\ell^{(0)} = \varepsilon_0$.

The density of states is obtained from the configurational average of the single-particle wave function. A convenient way to perform this average is the method of Matsubara and Toyozawa.[6] The Green's function results from the equation of motion,

$$G_{mj}(E) = G^{(0)}(E)\delta_{mj} + \sum_{k\ell} G^{(0)}(E)\delta_{mk}V_{k\ell}(E)G_{\ell j}(E) \, , \tag{6}$$

where $G^{(0)}(E) = (E - \epsilon_0)^{-1}$. Due to the overlapping of the $|\ell>$'s the hoping matrix becomes energy dependent,[7]

$$V_{ij}(E) = t_{ij} - (E - \epsilon_0)S_{ij} \ (1 - \delta_{ij}) \ , \tag{7}$$

where $S_{ij} = <i|j>$ and $t_{ij} = <i|V(\vec{r} - \vec{R}_i)|j>$.

In order to obtain the average Green's function it is necessary to calculate the Fourier transforms $S(\vec{k})$ and $V(\vec{k})$. A great deal of numerical calculation is avoided by noticing that

$$t(\vec{k}) = \int d^2R \ e^{-i\vec{k}\cdot\vec{R}} \ d^2r\phi(\vec{r})V(\vec{r})\phi(\vec{r} - \vec{R})$$

$$= \int d^2R \ e^{-i\vec{k}\cdot\vec{R}} \ d^2r\phi(\vec{r})\left[\epsilon_0 + \frac{\hbar^2\nabla^2}{2m^*}\right] \phi(\vec{r} - \vec{R}) \ . \tag{8}$$

After Fourier transforming $\phi(\vec{r})$, a simple expression for $V(\vec{k})$ comes out:

$$V(\vec{k}) = \left[2\epsilon_0 - E - \frac{\hbar^2k^2}{2m^*}\right] \phi^2(\vec{k}) \ . \tag{9}$$

If the overlap is neglected,[8]

$$V(\vec{k}) = t(\vec{k}) = \left[\epsilon_0 - \frac{\hbar^2k^2}{2m^*}\right]\phi^2(\vec{k}) \ . \tag{10}$$

Recently these calculations were performed and the 2-D density of states obtained both for the strictly bidimensional and unscreened model, i.e., z_0 and s zero,[9] and for unrestricted values of the parameters.[10] In the first case, where the overlap was also neglected, the discontinuity of the density of the states at the band edges was destroyed by disorder, but no band tailing has been observed. In the second case, overlapping effects were taken into account and the band tailing appears. It points to the higher values of energy in the intervals of s and z_0 considered, but this does not seem to be a

general characteristic of those bands. The calculated bandwidths for the parameter corresponding to the Si/SiO$_2$ (100) MOSFET are of the same order of magnitude of measured values.[10]

A surprising result comes out from some experiments, however: in some samples, the bandwidth of the impurity states decrease with the increase of Na$^+$ concentration. This is the opposite behaviour observed in doped semiconductors. The reason for this is expected to come from the screening mechanisms of the impurity potential. In fact, due to the band tailing of the first subband and its penetration by the impurity band, the density of conducting states may become dependent of the position of the Fermi level. Also, the process of screening by localized electrons is still an open question. So, a self-consistent treatment for the screening is requested at this point.

[1] See the review article of Prof. F. Stern in this Conference.

[2] See the review article of Prof. P. Stiles in this Conference.

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

[4] F. Stern and W.E. Howard, Phys. Rev. 163, 816 (1967).

[5] J. Serre and A. Gazhali, Phys. Rev. B28, 4704 (1983).

[6] T. Matsubara and Y. Toyozawa, Prog. Theor. Phys. 26, 739 (1961).

[7] Y. Ishida and F. Yonezawa, Prog. Theor. Phys. 49, 731 (1973).

[8] Eq. (10) is valid for any dimension. In 3-D, using the ground state wave function of the hydrogen atom for $\phi(\vec{r})$, and defining $V(k = q/a_0) = v(q).[a_0^3 \times Ry^\star]$, where a_0 is the effective Bohr radius, $v(q)$ becomes $-64\pi q^2/(1 + q^2)^4$ and $-64\pi\dfrac{(2 + q^2)}{(1 + q^2)^4}$, neglecting and taking into account the overlap, respectively. Curiously the calculations of Ref. 6 are performed using $v(q) = -64\pi/(1 + q^2)^3$.

[9] I.C. da Cunha Lima and A. Ferreira da Silva, Phys. Rev. B30, 4819 (1984).

[10] I.C. da Cunha Lima, A. Ferreira da Silva, P.S. Guimarães, L.F. Perondi and J.R. Senna, Phys. Rev. B (in press).