LOW-TEMPERATURE OPERATION

S.Selberherr

Technical University Vienna
Institut für Mikroelektronik
Gußhausstraße 27-29
A-1040 Wien
AUSTRIA

ABSTRACT: The state of the art in self consistent numerical low temperature modeling with particular emphasis on MOS devices is reviewed. The physical assumptions which are required to describe carrier transport at low ambient temperatures are discussed. Particular emphasis is put on the models for space charge (impurity freeze-out), carrier mobility (temperature dependence of scattering mechanisms) and carrier generation-recombination (impact ionization). The differences with regard to the numerical methods required for the solution of low temperature models compared to room temperature models are explained. Typical results obtained with the simulator MINIMOS are presented. These include comparisons of short channel effects and hot-electron phenomena like energy relaxation and avalanche breakdown at 77K and 300K ambient temperature.

I. INTRODUCTION

Device Modeling based on the self-consistent solution of fundamental semiconductor equations dates back to the famous work of Gummel in 1964 [33]. However, the first application of this rigorous style of modeling for problems at low ambient temperature (usually liquid nitrogen temperature) has first been carried out by Gaensslen et al. about twelve years later in 1976 [29]. The main reason for this delay cannot only be seen in the lesser demands for low temperature simulation. Not only supercomputers have been made for operation at liquid nitrogen temperature, like the ETA-10 [25], but also microprocessors, cf. [19]. The primary reason for the fairly poor status in fully numerical low temperature device modeling stems from considerably increased difficulties regarding physical assumptions and implementation of the numerical solution.

This manuscript is an enhanced version of reference [60].

II. PHYSICAL ASPECTS

a) Basic Equations

The model for hot carrier transport used in any numerical device simulation is based on the well known fundamental semiconductor equations (1)-(5). There are ongoing arguments in the scientific community whether these equations are adequate to describe transport in submicron devices. Particularly the current relations (4) and (5) which are the most complex equations out of the set of the basic semiconductor device equations undergo strong criticism in view of, for instance, ballistic transport [37], [51]. Their derivation from more fundamental physical principles is indeed not at all straightforward. They appear therefore with all sorts of slight variations in the specialized literature and a vast number of papers has been published where some of their subtleties are dealt with. The interested reader is referred to, e.g., [9], [13], [28], [58]. Anyway, recent investigations on ultra short MOSFET's [52] do not give evidence that it is necessary to waive these well established basic equations for silicon devices down to feature sizes in the order of $0,1$ microns [62].

$$\operatorname{div}(\varepsilon \cdot \operatorname{grad} \psi) = -\rho \tag{1}$$

$$\operatorname{div} \vec{J}_n - q \cdot \frac{\partial n}{\partial t} = q \cdot R \tag{2}$$

$$\operatorname{div} \vec{J}_p + q \cdot \frac{\partial p}{\partial t} = -q \cdot R \tag{3}$$

$$\vec{J}_n = q \cdot \mu_n \cdot n \cdot \left(\vec{E} + \frac{1}{n} \cdot \operatorname{grad} \left(n \cdot \frac{k \cdot T_n}{q} \right) \right) \tag{4}$$

$$\vec{J}_p = q \cdot \mu_p \cdot p \cdot \left(\vec{E} - \frac{1}{p} \cdot \operatorname{grad} \left(p \cdot \frac{k \cdot T_p}{q} \right) \right) \tag{5}$$

These equations include a set of parameters which have to be appropriately modeled in order to correctly describe the various transport phenomena qualitatively and quantitatively.

b) Modeling Space Charge

Poisson's equation (1) requires a model for the space charge ρ which makes use of only the dependent variables ψ, n, p and material properties. The well established approach

for this model is to sum up the concentrations with the adequate charge sign multiplied with the elementary charge (6).

$$\rho = q \cdot (p - n + N_D^+ - N_A^-) \tag{6}$$

Here the first difference between room temperature and low temperature simulation becomes apparent. The doping concentration $N_D^+ - N_A^-$ is usually assumed to be fully ionized at room temperature which intuitively does not hold for low temperature analysis. The classical way to describe partial ionization is based on the formulae (7).

$$N_D^+ = \frac{N_D}{1 + 2 \cdot \exp\left(\frac{E_{fn} - E_D}{k \cdot T}\right)}, \qquad N_A^- = \frac{N_A}{1 + 4 \cdot \exp\left(\frac{E_A - E_{fp}}{k \cdot T}\right)} \tag{7}$$

E_D and E_A are the ionization energies of the respective donor and acceptor dopant. Typical values for $E_c - E_D$ and $E_A - E_v$ for the most common dopants in silicon are: 0.054eV for arsenic, 0.045eV for phosphorus, 0.039eV for antimony and 0.045eV for boron. A quite complete list can be found in [67]. These ionization energies are recommended to be modeled doping dependent in [17], however, it seems not to be important for MOSFET's regarding my experience. Note, that only energy differences can be given (E_c and E_v are the conduction band and the valence band energy, respectively). Next the Fermi levels E_{fn} and E_{fp} have to be appropriately related to the dependent variables by making use of Fermi statistics.

$$n = N_c \cdot \frac{2}{\sqrt{\pi}} \cdot F_{1/2}\left(\frac{E_{fn} - E_c}{k \cdot T}\right), \qquad p = N_v \cdot \frac{2}{\sqrt{\pi}} \cdot F_{1/2}\left(\frac{E_v - E_{fp}}{k \cdot T}\right) \tag{8}$$

N_c and N_v are the density of states in the conduction band and the valence band, respectively. The classical formulae for the density of states are given by (9).

$$N_c = 2 \cdot \left(\frac{2 \cdot \pi \cdot k \cdot T \cdot m_n^*}{h^2}\right)^{3/2}, \qquad N_v = 2 \cdot \left(\frac{2 \cdot \pi \cdot k \cdot T \cdot m_p^*}{h^2}\right)^{3/2} \tag{9}$$

$F_{1/2}(x)$ is the Fermi function of order 1/2 which is defined by (10).

$$F_{1/2}(x) = \int_0^\infty \frac{\sqrt{y}}{1 + e^{y-x}} \cdot dy \tag{10}$$

The parameters m_n^* and m_p^* which are the effective masses for electrons and holes have now to be modeled to be able to evaluate the formulae for the density of states (9). The probably most elaborate models which are fits to experimental values date back to Gaensslen et al. [29], [31].

$$m_n^* = m_o \cdot \left(1,045 + 4,5 \cdot 10^{-4} \cdot \left(\frac{T}{\mathrm{K}}\right)\right) \tag{11}$$

$$m_p^* = m_o \cdot \left(0,523 + 1,4 \cdot 10^{-3} \cdot \left(\frac{T}{\mathrm{K}}\right) - 1,48 \cdot 10^{-6} \cdot \left(\frac{T}{\mathrm{K}}\right)^2\right) \tag{12}$$

These fitting expressions are claimed to be valid over the range 50-350K. A graphical representation is given in Fig.1.

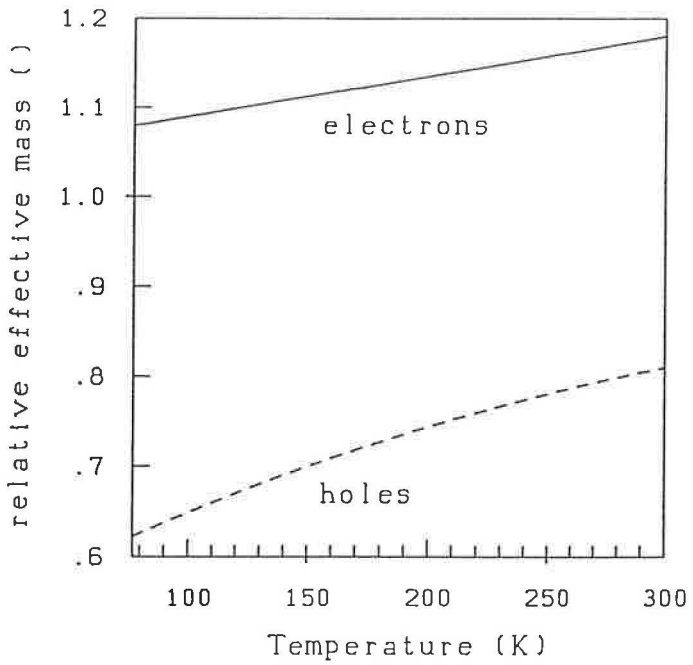

Fig.1: Relative Effective Masses Versus Temperature in Silicon

It is worthwhile to note that the ratio of the density of states depends only on the ratio of the effective masses.

$$\frac{N_c}{N_v} = \left(\frac{m_n^*}{m_p^*}\right)^{3/2} \tag{13}$$

By means of some simple algebraic manipulation with the expressions for the carrier concentrations (8) we obtain:

$$\frac{E_{fn} - E_D}{k \cdot T} = G_{1/2}\left(\frac{n}{N_c}\right) + \frac{E_c - E_D}{k \cdot T} \tag{14}$$

$$\frac{E_A - E_{fp}}{k \cdot T} = G_{1/2}\left(\frac{p}{N_v}\right) + \frac{E_A - E_v}{k \cdot T} \tag{15}$$

$G_{1/2}(x)$ is the inverse Fermi function of order 1/2 defined with (16).

$$G_{1/2}\left(\frac{2}{\sqrt{\pi}} \cdot F_{1/2}(x)\right) = x \tag{16}$$

It is now possible by evaluating the expressions for the density of states (9) with the fits to the effective masses (11) and (12) to compute numerical values for the ionized impurity concentrations (7) using only the carrier concentrations which are the dependent variables in the basic equations. However, comparisons to experiment indicate that it is better to compute the density of states from relation (13) and a fit to the intrinsic carrier concentration (17).

$$n_i = \sqrt{N_c \cdot N_v} \cdot \exp\left(-\frac{E_g}{2 \cdot k \cdot T}\right) \tag{17}$$

E_g is the band gap $E_c - E_v$ which can be modeled temperature dependent with the fit provided by Gaensslen et al. [29], [31]. Note that most publications which present equation (18) contain a typographical error. The linear temperature coefficient for E_g below 170K is $1,059 \cdot 10^{-5}$eV and not as usually found $1,059 \cdot 10^{-6}$eV [40]. A graphical representation of (18) is given in Fig.2.

$$E_g = \begin{cases} 1,17\text{eV} + 1,059 \cdot 10^{-5}\text{eV} \cdot \left(\frac{T}{\text{K}}\right) - 6,05 \cdot 10^{-7}\text{eV} \cdot \left(\frac{T}{\text{K}}\right)^2 & T \leq 170\text{K} \\ 1,1785\text{eV} - 9,025 \cdot 10^{-5}\text{eV} \cdot \left(\frac{T}{\text{K}}\right) - 3,05 \cdot 10^{-7}\text{eV} \cdot \left(\frac{T}{\text{K}}\right)^2 & T > 170\text{K} \end{cases} \tag{18}$$

The prefactor in (17) can be fitted to experiments by (19). With these data one obtains an intrinsic carrier concentration as given with Fig.3. Note the well known extreme dynamic range of thirty orders of magnitude which causes severe numerical implementation problems for device simulation programs.

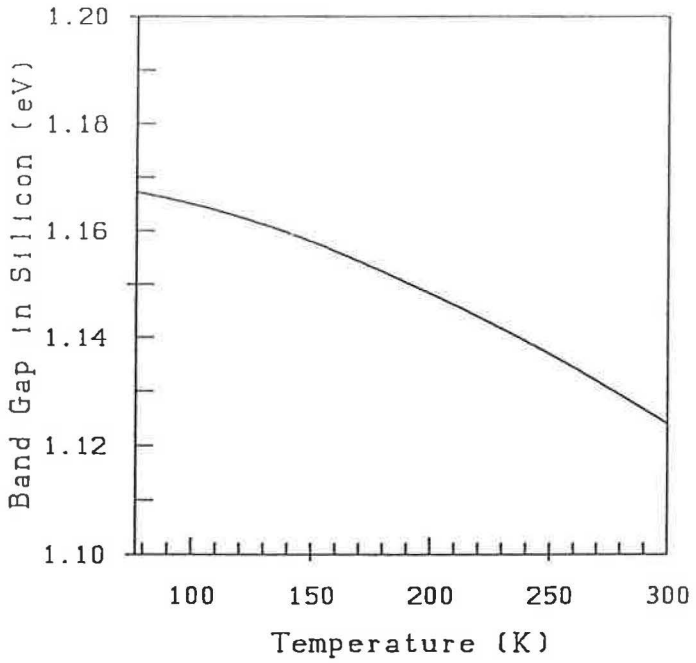

Fig.2: Energy Band Gap Versus Temperature in Silicon

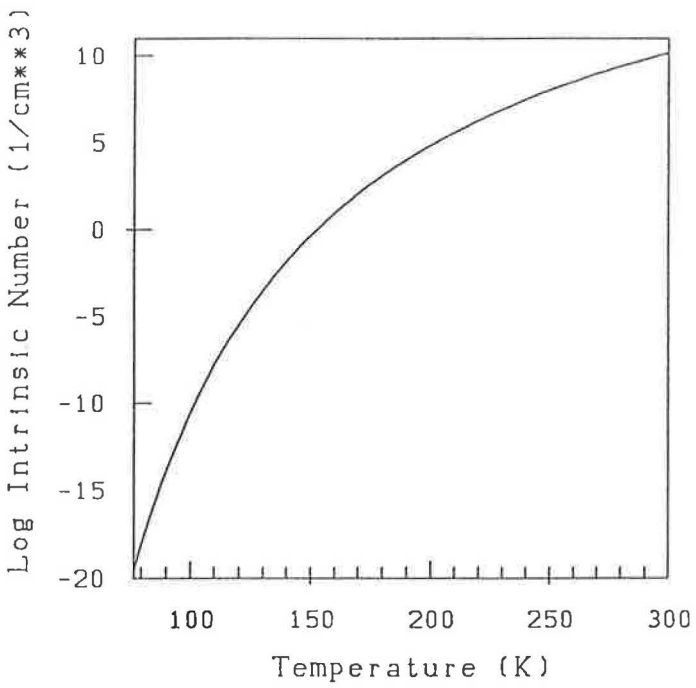

Fig.3: Intrinsic Carrier Concentration Versus Temperature in Silicon

$$\sqrt{N_c \cdot N_v} = \exp\left(45,13 + 0,75 \cdot \ln\left(\frac{m_n^*}{m_o} \cdot \frac{m_p^*}{m_o} \cdot \left(\frac{T}{300\mathrm{K}}\right)^2\right)\right) \mathrm{cm}^{-3} \qquad (19)$$

With (13) and (19) it is now straightforward to compute the numerical values for the density of states. At room temperature we have $N_c = 5,1 \cdot 10^{19}\mathrm{cm}^{-3}$ and $N_v = 2,9 \cdot 10^{19}\mathrm{cm}^{-3}$; at liquid nitrogen temperature we obtain $N_c = 5,8 \cdot 10^{18}\mathrm{cm}^{-3}$ and $N_v = 2,5 \cdot 10^{18}\mathrm{cm}^{-3}$.

It should be noted that equations (7) are only valid for moderate impurity concentrations. For heavy doping the assumption of a localized ionization energy does definitely not hold. Instead an impurity band is formed which may merge with the respective band edge, e.g. [38], [55]. Just modeling a temperature dependence of the ionization energies will not account adequately for the underlying physics in this case. It appears to be appropriate to assume total ionization for concentrations above some threshold value and to account for a suitable functional transition between the classical formulae (7) and total ionization. All concepts to tackle this problem which have come to my attention so far, however, make use of a very simplistic, not to say alchemical, approach. In MINIMOS these threshold concentrations are $10^{18}cm^{-3}$ and $10^{19}cm^{-3}$, respectively. For a dopant concentration below the lower threshold concentration the respective of equation (7) is used to compute the ionized dopant concentration; for a dopant concentration above the higher threshold concentration complete ionization is assumed. In the intervall between the two threshold concentrations linear interpolation is performed on the fraction of ionization to obtain a smooth transistion between partial and complete ionization. Anyway, it should be noted that freeze-out is of major importance only for depletion mode devices and devices with a partially compensated channel doping [32].

In view of this dilemma with heavy doping one may for many applications well use asymptotic approximations for the Fermi function (10) and its inverse (16) (see [57]). This is not in contradiction to the partial ionization model given with (6)-(19).

c) Modeling Carrier Mobilities

The next set of physical parameters to be considered carefully for low temperature simulation consists of the carrier mobilities μ_n and μ_p in (4) and (5). The models for the carrier mobilities have to take into account a great variety of scattering mechanisms the most basic one of which is lattice scattering. The lattice mobility in pure silicon can be fitted with simple power laws.

$$\mu_n^L = 1430\frac{\mathrm{cm}^2}{\mathrm{Vs}} \cdot \left(\frac{T}{300\mathrm{K}}\right)^{-2}, \qquad \mu_p^L = 460\frac{\mathrm{cm}^2}{\mathrm{Vs}} \cdot \left(\frac{T}{300\mathrm{K}}\right)^{-2,18} \qquad (20)$$

The expressions (20) fit well experimental data of [2], [15] and [47].

The next effect to be considered is ionized impurity scattering. The best established procedure for this task is to take the functional form (21) of the fit provided by Caughey and Thomas [16] and use temperature dependent coefficients.

$$\mu_{n,p}^{LI} = \mu_{n,p}^{min} + \frac{\mu_{n,p}^{L} - \mu_{n,p}^{min}}{1 + \left(\frac{CI}{C_{n,p}^{ref}}\right)^{\alpha_{n,p}}} \tag{21}$$

$$\mu_n^{min} = \begin{cases} 80\frac{\mathrm{cm}^2}{\mathrm{Vs}} \cdot \left(\frac{T}{300\mathrm{K}}\right)^{-0,45} & T \geq 200\mathrm{K} \\ 80\frac{\mathrm{cm}^2}{\mathrm{Vs}} \left(\frac{200\mathrm{K}}{300\mathrm{K}}\right)^{-0,45} \cdot \left(\frac{T}{200\mathrm{K}}\right)^{-0,15} & T < 200\mathrm{K} \end{cases} \tag{22}$$

$$\mu_p^{min} = \begin{cases} 45\frac{\mathrm{cm}^2}{\mathrm{Vs}} \cdot \left(\frac{T}{300\mathrm{K}}\right)^{-0,45} & T \geq 200\mathrm{K} \\ 45\frac{\mathrm{cm}^2}{\mathrm{Vs}} \left(\frac{200\mathrm{K}}{300\mathrm{K}}\right)^{-0,45} \cdot \left(\frac{T}{200\mathrm{K}}\right)^{-0,15} & T < 200\mathrm{K} \end{cases} \tag{23}$$

$$C_n^{ref} = 1,12 \cdot 10^{17}\mathrm{cm}^{-3} \cdot \left(\frac{T}{300\mathrm{K}}\right)^{3,2}, \qquad C_p^{ref} = 2,23 \cdot 10^{17}\mathrm{cm}^{-3} \cdot \left(\frac{T}{300\mathrm{K}}\right)^{3,2} \tag{24}$$

$$\alpha_{n,p} = 0,72 \cdot \left(\frac{T}{300\mathrm{K}}\right)^{0,065} \tag{25}$$

The fits (22)-(25) are from [36]. Similar data have been provided in [6] and [24].

In view of partial ionization one should consider neutral impurity scattering [57]. However, in view of the uncertainty of the quantitative values for ionized impurity scattering it seems not to be worthwhile to introduce another scattering mechanism with additional fitting parameters. Furthermore, partial ionization appears to be a second order effect even at liquid nitrogen temperature. It seems therefore justified to include partial ionization only in the space charge model and not in the carrier mobilities.

Surface scattering is modeled with an expression suggested by Seavey [54].

$$\mu_{n,p}^{LIS} = \frac{\mu_{n,p}^{ref} + (\mu_{n,p}^{LI} - \mu_{n,p}^{ref}) \cdot (1 - F(y))}{1 + F(y) \cdot \left(\frac{S_{n,p}}{S_{n,p}^{ref}}\right)^{\gamma_{n,p}}} \tag{26}$$

$$\mu_n^{ref} = 638\frac{\mathrm{cm}^2}{\mathrm{Vs}} \cdot \left(\frac{T}{300\mathrm{K}}\right)^{-1,19}, \qquad \mu_p^{ref} = 240\frac{\mathrm{cm}^2}{\mathrm{Vs}} \cdot \left(\frac{T}{300\mathrm{K}}\right)^{-1,09} \tag{27}$$

with:

$$F(y) = \frac{2 \cdot \exp\left(-\left(\frac{y}{y^{ref}}\right)^2\right)}{1 + \exp\left(-2 \cdot \left(\frac{y}{y^{ref}}\right)^2\right)} \tag{28}$$

$$S_n = \max\,(0,\, \frac{\partial\psi}{\partial y}), \qquad S_p = \max\,(0, -\frac{\partial\psi}{\partial y}) \tag{29}$$

S_n^{ref} is assumed to be $7 \cdot 10^5 \frac{\mathrm{V}}{\mathrm{cm}}$; S_p^{ref} is $2,7 \cdot 10^5 \frac{\mathrm{V}}{\mathrm{cm}}$; γ_n is 1,69; γ_p is 1,0 and y^{ref} is 10nm.

The formulae for surface scattering are definitely not the ultimate expressions. They just fit quite reasonably experimental observations. Other approaches with the same claim can be found in, e.g., [7], [39], [49]. A u-shaped mobility behavior as found in [8], [10] has not been synthesized because I believe in a different origin than surface scattering for this experimental observation. It should however be noted that soft turn-on at liquid nitrogen temperature has been successfully simulated with a u-shaped mobility expression [27].

Velocity saturation is modeled with formulae (30). These are again fits to experimental data with, however, a theoretical background considering their functional form [2], [42], [43].

$$\mu_n^{LISE} = \frac{2 \cdot \mu_n^{LIS}}{1 + \sqrt{1 + \left(\frac{2 \cdot \mu_n^{LIS} \cdot E_n}{v_n^{sat}}\right)^2}}, \qquad \mu_p^{LISE} = \frac{\mu_p^{LIS}}{1 + \frac{\mu_p^{LIS} \cdot E_p}{v_p^{sat}}} \tag{30}$$

E_n and E_p are the effective driving forces given by (31). Their derivation can be found in [34].

$$E_n = |\mathrm{grad}\ \psi - \frac{1}{n} \cdot \mathrm{grad}\ (Ut_n \cdot n)|, \qquad E_p = |\mathrm{grad}\ \psi + \frac{1}{p} \cdot \mathrm{grad}\ (Ut_p \cdot p)| \tag{31}$$

The saturation velocities are given in the following.

$$v_n^{sat} = 1,45 \cdot 10^7 \frac{\text{cm}}{\text{s}} \cdot \sqrt{\tanh\left(\frac{155\text{K}}{T}\right)}, \qquad v_p^{sat} = 9,05 \cdot 10^6 \frac{\text{cm}}{\text{s}} \cdot \sqrt{\tanh\left(\frac{312\text{K}}{T}\right)} \quad (32)$$

The functional form of these fits is after [2]; the experimental data matched are from [2], [14], [15], [22]. An eventual dependence on the crystallographic orientation which one would deduce from [3], [5], [45] is presently not taken into account.

A review of the evolution of this mobility model can be found in [61].

d) Modeling Carrier Temperatures

To describe carrier heating properly one has to account for local carrier temperatures $T_{n,p}$ in the current relations (4) and (5). This can be achieved by either solving energy conservation equations self consistently with the basic transport equations, or by using a model obtained by series expansions of the solution to the energy conservation equations [34]. I believe that the latter is sufficient for silicon devices. For the electronic voltages we have (33) as an approximation. Confirming theoretical investigations can be found in [1].

$$Ut_{n,p} = \frac{k \cdot T_{n,p}}{q} = Ut_o + \frac{2}{3} \cdot \tau_{n,p}^{\epsilon} \cdot \left(v_{n,p}^{sat}\right)^2 \cdot \left(\frac{1}{\mu_{n,p}^{LISE}} - \frac{1}{\mu_{n,p}^{LIS}}\right) \quad (33)$$

The energy relaxation times $\tau_{n,p}^{\epsilon}$ are in the order of $0,5$picoseconds and just weakly temperature dependent [11]. They should however be modeled as functions of the local doping concentration as motivated by the following reasoning. The product of carrier mobility times electronic voltage which symbolizes a diffusion coefficient must be a decreasing function with increasing carrier voltage (see also [11]). Its maximum is attained at thermal equillibrium. Relation (34) must therefore hold.

$$\mu_{n,p}^{LISE} \cdot Ut_{n,p} \leq \mu_{n,p}^{LIS} \cdot Ut_o \quad (34)$$

Note that models for carrier diffusion coefficients are not required in the basic current relations (4), (5).

Substituting (33) into (34) and rearranging terms one obtains relation (35) for the energy relaxation times.

$$\tau_{n,p}^{\epsilon} \leq \frac{3}{2} \cdot Ut_o \cdot \frac{\mu_{n,p}^{LIS}}{\left(v_{n,p}^{sat}\right)^2} \quad (35)$$

In MINIMOS the energy relaxation times are modeled on the basis of (35) with a fudge factor γ in the range $[0,1]$ and a default value of $0,8$.

$$\tau_{n,p}^{\epsilon} = \gamma \cdot \frac{3}{2} \cdot Ut_o \cdot \frac{\mu_{n,p}^{LIS}}{\left(v_{n,p}^{sat}\right)^2} \tag{36}$$

For vanishing doping one obtains the maximal energy relaxation times which are at 300K $\tau_n^{\epsilon} = 4,44{\cdot}10^{-13}$s, $\tau_p^{\epsilon} = 2,24{\cdot}10^{-13}$s and at liquid nitrogen temperature $\tau_n^{\epsilon} = 8,82{\cdot}10^{-13}$s, $\tau_p^{\epsilon} = 8,68 \cdot 10^{-13}$s.

e) Modeling Carrier Generation/Recombination

Thermal generation/recombination at 77K can be modeled the same way as at 300K [57]. A comment should be made on the model for the impact ionization rate which has to be supplied for the continuity equations (2) and (3). It still seems, though under heavy dispute in the scientific community, that the old Chynoweth formulation (37) of impact ionization can be used quite satisfactorily for device simulation.

$$R^{II} = -\alpha_n \cdot \frac{|\vec{J}_n|}{q} - \alpha_p \cdot \frac{|\vec{J}_p|}{q} \tag{37}$$

with:

$$\alpha_{n,p} = \alpha_{n,p}^{\infty} \cdot \exp\left(-\frac{\beta_{n,p}}{E}\right) \tag{38}$$

The coefficients of (38) can be modeled temperature dependent by (39) and (40) to fit experimental data [20], [23], [50]. It should be noted that there is some lack of data for liquid nitrogen temperature, cf. [66]. However it seems that this impact ionization model is probably somewhat too pessimistic for a proper quantitative prediction of substrate currents as already stated in [46], [59].

$$\begin{aligned} \alpha_n^{\infty} &= 7 \cdot 10^5 \mathrm{cm}^{-1} \cdot \left(0,57 + 0,43 \cdot \left(\frac{T}{300\mathrm{K}}\right)^2\right) \\ \alpha_p^{\infty} &= 1,58 \cdot 10^6 \mathrm{cm}^{-1} \cdot \left(0,58 + 0,42 \cdot \left(\frac{T}{300\mathrm{K}}\right)^2\right) \end{aligned} \tag{39}$$

$$\beta_n = 1,23 \cdot 10^6 \frac{\mathrm{V}}{\mathrm{cm}} \cdot \left(0,625 + 0,375 \cdot \left(\frac{T}{300\mathrm{K}}\right)\right)$$

$$\beta_p = 2,04 \cdot 10^6 \frac{\mathrm{V}}{\mathrm{cm}} \cdot \left(0,67 + 0,33 \cdot \left(\frac{T}{300\mathrm{K}}\right)\right) \tag{40}$$

The Auger coefficients for the model of Auger recombination (41) can also be made weakly temperature dependent with (42). The fit has been made to the data of [26].

$$R^{AU} = (C_{cn} \cdot n + C_{cp} \cdot p) \cdot (n \cdot p - {n_i}^2) \tag{41}$$

$$C_{cn} = 2,8 \cdot 10^{-31} \frac{\mathrm{cm}^6}{\mathrm{s}} \cdot \left(\frac{T}{300\mathrm{K}}\right)^{0,14}, \qquad C_{cp} = 9,9 \cdot 10^{-32} \frac{\mathrm{cm}^6}{\mathrm{s}} \cdot \left(\frac{T}{300\mathrm{K}}\right)^{0,2} \tag{42}$$

III. NUMERICAL ASPECTS

Almost none of the many device simulation programs which have proven their usefulness for room temperature simulations can be directly applied to low temperature applications. The primary reason for this is the scaling of carrier concentrations with all thereby induced consequences. Briefly sketched: De Mari recommended in an early paper [21] to scale the intrinsic carrier concentration to unity, which contributes to change the basic semiconductor equations into a dimensionless form very convenient for computer implementation. Due to the fact that the intrinsic carrier concentration at liquid nitrogen temperature is in the order of $10^{-20}\mathrm{cm}^{-3}$ (cf.Fig.3), it is obviously not applicable for scaling in this case, since for instance an impurity concentration of $10^{20}\mathrm{cm}^{-3}$ would then be scaled to 10^{40}. To scale the maximum impurity concentration to unity as recommended in the elaborate mathematical book of Markowich [48] is also not feasible, since the scaled intrinsic concentration would be in the order of 10^{-40}. A way out of this dilemma is to use the concentration C_s defined by (43) for scaling.

$$C_s = \sqrt{4,8 \cdot 10^{22}\mathrm{cm}^{-3} \cdot n_i} \tag{43}$$

The magic concentration in (43) is the number of silicon atoms per cubic centimeter. It serves as an absolute upper limit for the maximum possible concentration of any type. At 300K temperature $C_s = 2,58 \cdot 10^{16}\mathrm{cm}^{-3}$, at 77K temperature $C_s = 3,51 \cdot 10^{1}\mathrm{cm}^{-3}$. This scaling equilibrates the relevant concentrations and thus is optimal for avoiding fatal floating point underflow or overflow exceptions.

A further detail to be considered for numerical implementation is the appropriate approximation of the inverse Fermi function $G_{1/2}(x)$. A convenient fit is given by (44).

$$G_{1/2}(x) = \frac{\ln(x)}{1-x^2} + \frac{\left(\frac{3\cdot\sqrt{\pi}\cdot x}{4}\right)^{2/3}}{1 + \frac{1}{\left(0,24 + 1,08\cdot\left(\frac{3\cdot\sqrt{\pi}\cdot x}{4}\right)^{2/3}\right)^2}} \tag{44}$$

The first term in (44) has to be replaced by a truncated series expansion if the argument x is in the vicinity of 1.

$$\frac{\ln(x)}{1-x^2} = \frac{x-2}{2} + O((x-1)^2) \tag{45}$$

A review about approximations to Fermi functions and their inverse functions can be found in [12], [18], [57].

IV. A GLIMPSE OF RESULTS

Results of investigations about submicron n-channel enhancement mode MOSFET's at room and liquid nitrogen ambient temperature are presented. A single drain technology designed for well functioning devices at room temperature operation with 3/4 gate length (0,39 micron metallurgical channel length) has been analyzed where the gate length has been shrunk for liquid nitrogen temperature operation to 0,51 micron (0,15 micron metallurgical channel length). The gate oxide thickness is 9nm and the work function of the donor doped gate polysilicon is -570mV. A window of the critical drain profile corner is depicted in Fig.4.

The actual analysis has been carried out with MINIMOS 4 which is a further development of the MINIMOS program [34], [53], [56] to include also quantitative capabilities for low temperature simulation.

The threshold voltage versus gate length L is shown in Fig.5 for this particular technology. The long channel threshold voltage is about 580mV at 77K and about 340mV at room temperature. Threshold voltage is here defined as the gate voltage required to achieve a drain current of $0,1\mu A \cdot W/L$ at a given drain bias. No substrate bias has been applied for these investigations. It appears to be a matter of taste whether to use the gate length L or the metallurgical channel length for this kind of investigation. However, device designers usually prefer the gate length.

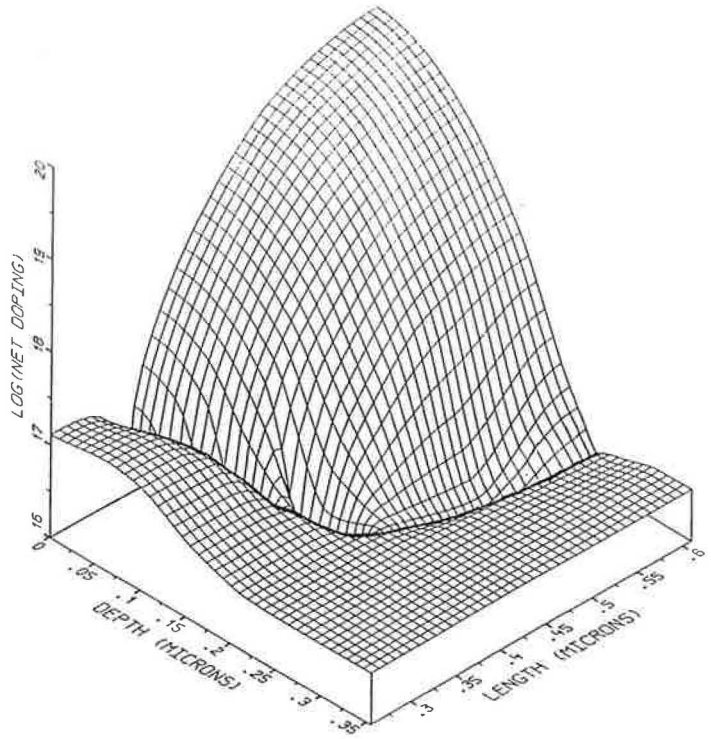

Fig.4: Detail of Net Impurity Profile

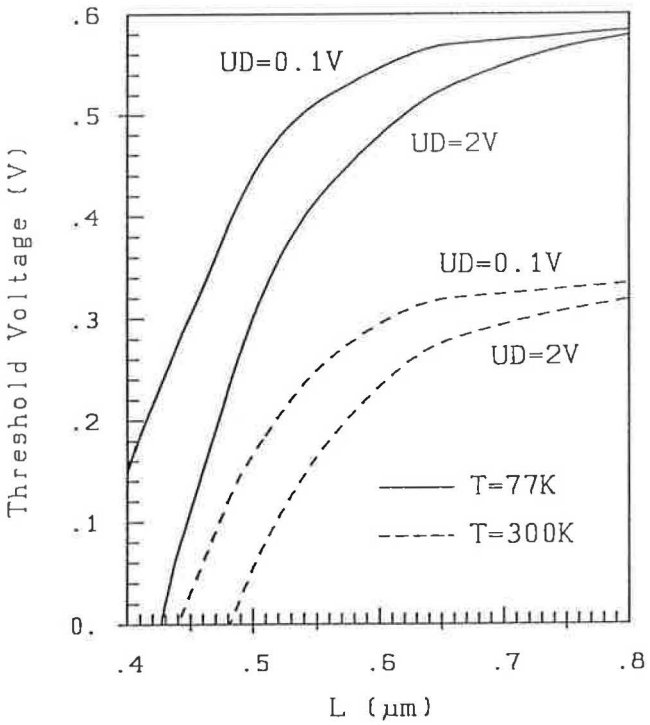

Fig.5: Threshold Voltage Versus Gate Length

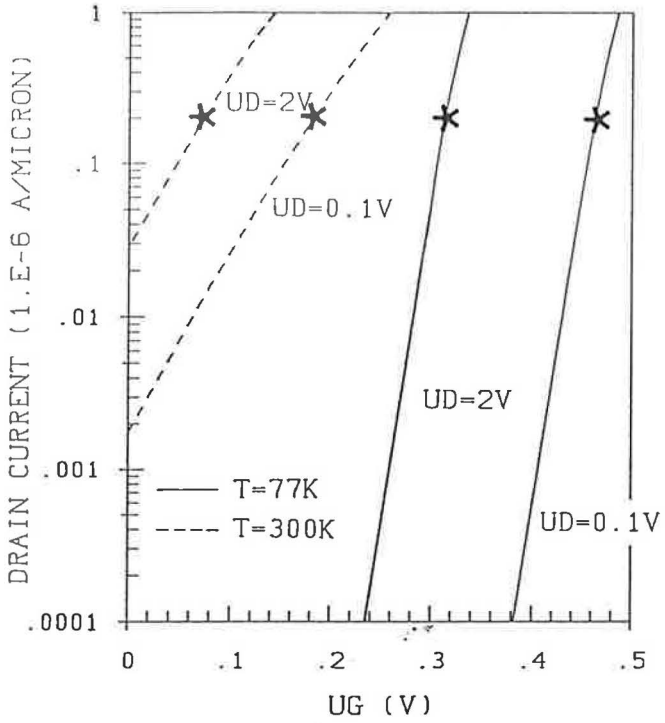

Fig.6: Simulated Subthreshold Characteristics
The asterisks indicate the respective threshold voltages

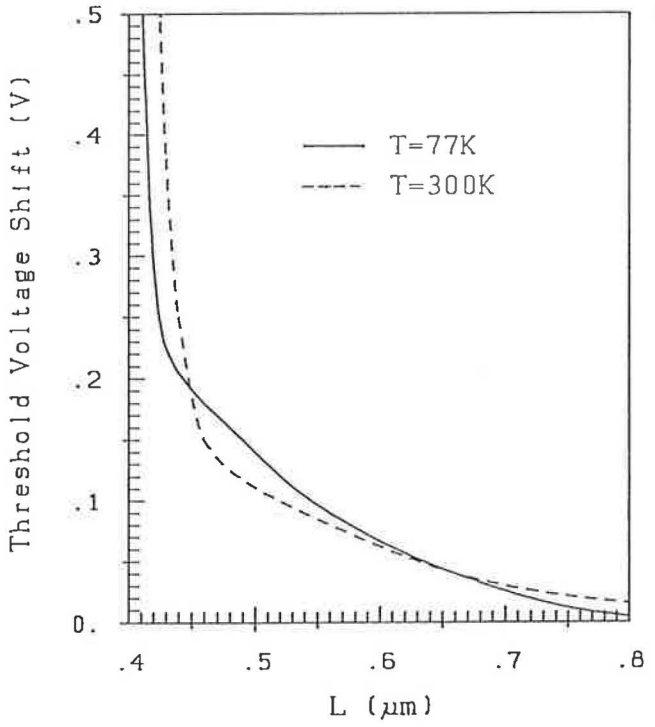

Fig.7: Drain Bias Induced Threshold Voltage Shift Versus Gate Length

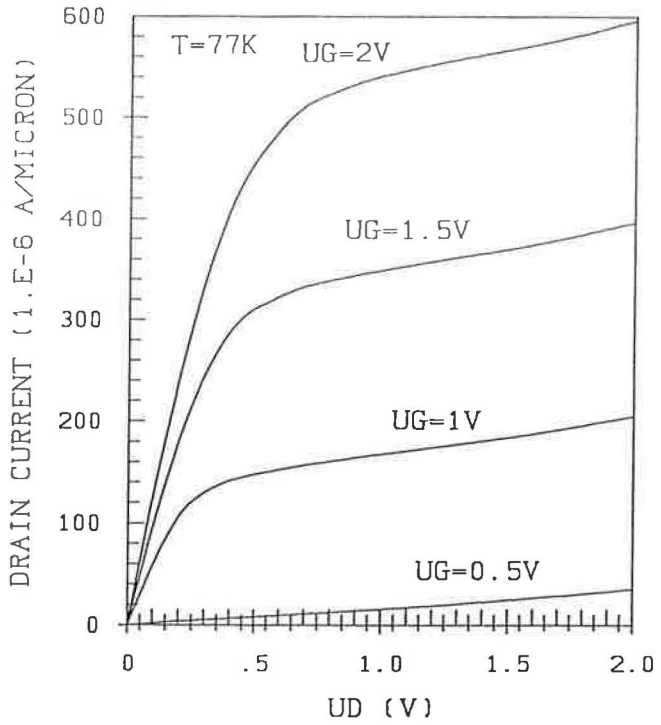

Fig.8: Simulated Output Characteristics at 77k

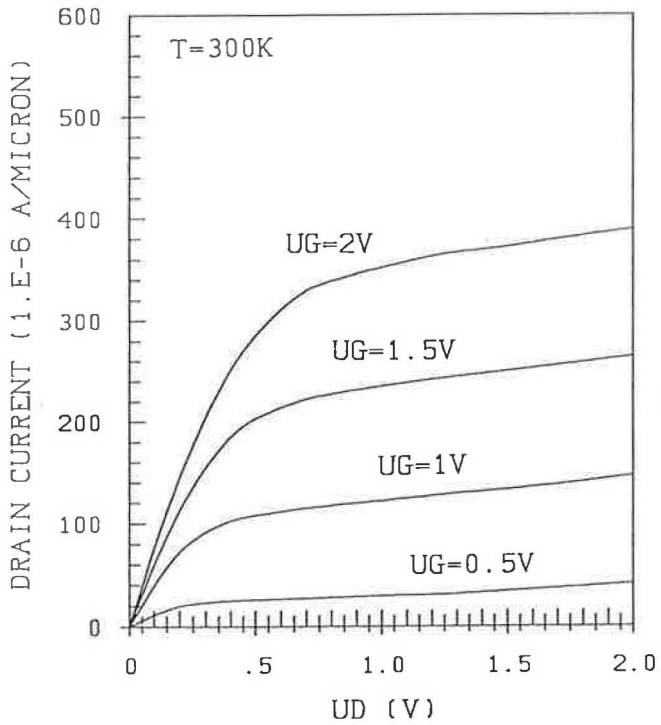

Fig.9: Simulated Output Characteristics at 300K

Fig.6 shows the simulated subthreshold characteristics for two different drain biases (UD=0,1V, UD=2V, UB=0V) at room and liquid nitrogen temperature. The subthreshold slope is obviously much steeper at liquid nitrogen temperature with about 25mV/decade compared to 95mV/decade at room temperature. It is interesting that the improvement of the slope is almost as good as 3,9 the ratio of 300K/77K [41]. Furthermore, the shift of the subthreshold characteristics between low and high drain bias which should be primarily due to drain induced surface barrier lowering is about 50mV larger at 77K temperature compared to the room temperature shift. To have a larger influence of drain induced barrier lowering at lower temperature is in contradiction with the sound results of [69]. The observed phenomenon must therefore be of different origin. In order to get insight into this effect the drain bias induced threshold voltage shift has been computed as a function of gate length. Fig.7 shows this shift of the threshold voltage between UD=0,1V and UD=2V for this technology. One can nicely see that the functional behavior is caused by two overlapping phenomena with a knee at a gate length of $0,43\mu m$ at 77K and $0,46\mu m$ at 300K. Detailed investigations have brought up several interacting causes. One is partial freeze-out of acceptors in the bulk below the channel which leads to an increase of built-in potential and thus to increasing depletion widths with decreasing temperature [44], [68]. This reasoning is partially confirmed in [30]. The second cause is the formation of a sort of parasitic channel by impact ionization which has also been reported in [53].

Fig.8 and Fig.9 show the simulated output characteristics for four different gate biases. If we take current output for the same gate drive as measure of device quality, the low temperature operation resulted in approximately 50% improvement compared to room temperature operation. Similar results have been experimentally obtained (cf.[64]). This improvement decreases with shrinking channel length as observed in [52].

In the following a few results about the distribution of the various physical quantities in the interior of the device will be presented. The off-state at UG=0V, UD=2V is depicted with the electron concentration given by Fig.10 and Fig.11 at liquid nitrogen temperature and room temperature, respectively. It is easily visible that the device is not satisfactorily off at room temperature. The channel is perfectly depleted at 77K.

The on-state is documented with a bias of UG=2V, UD=2V. Fig.12 and Fig.13 show again the electron concentration at liquid nitrogen temperature and room temperature, respectively. One can nicely observe the inversion layer which is much steeper for the low temperature simulation. Furthermore, one can see that there are considerably more electrons generated by impact ionization close to the drain.

The impact ionization rates are shown in Fig.14 and Fig.15. The peak concentration which occurs in both cases at the surface is almost two orders of magnitude higher for low temperature operation. The substrate current to drain current ratio is increased by a factor of 5.2 which is fairly high for n-channel devices [4].

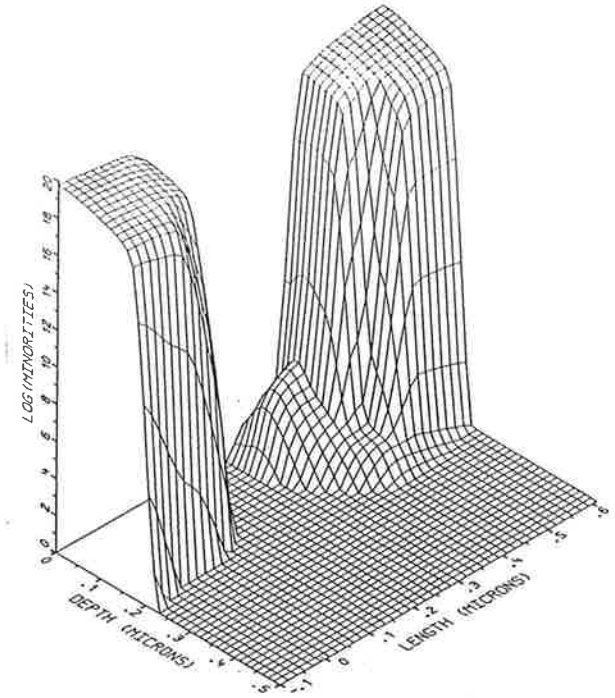

Fig.10: Electron Concentration at 77K (UG=0V, UD=2V)

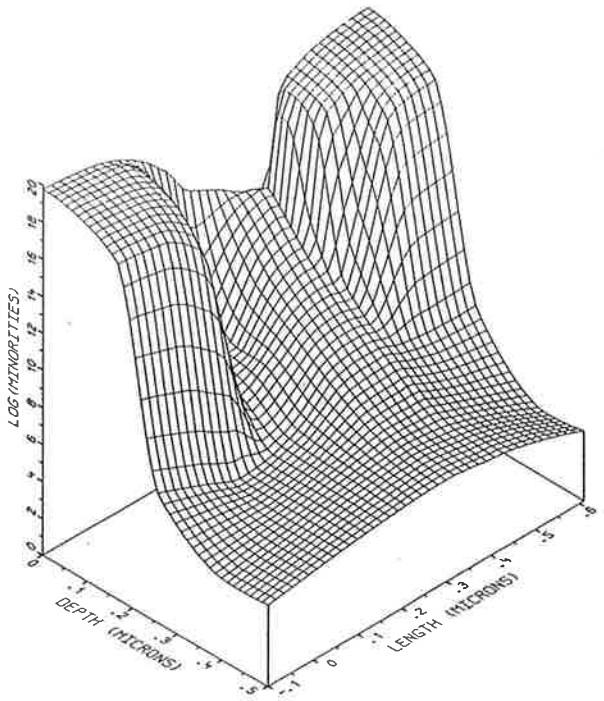

Fig.11: Electron Concentration at 300K (UG=0V, UD=2V)

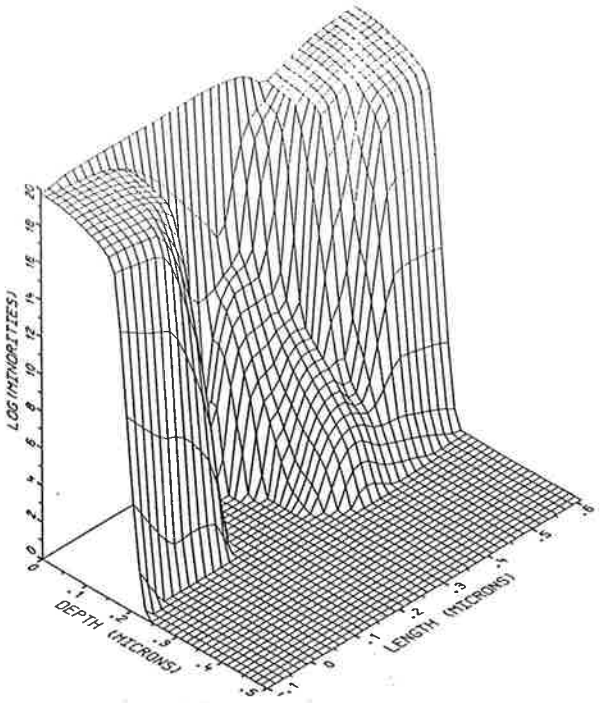

Fig.12: Electron Concentration at 77K (UG=2V, UD=2V)

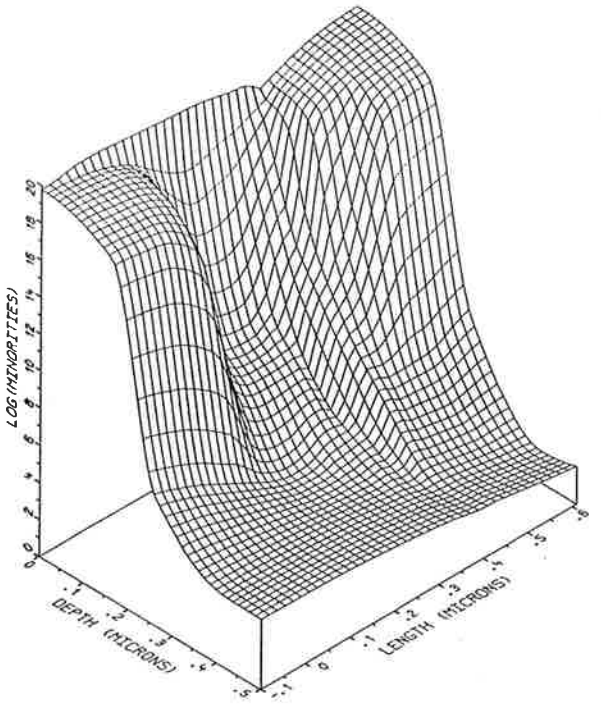

Fig.13: Electron Concentration at 300K (UG=2V, UD=2V)

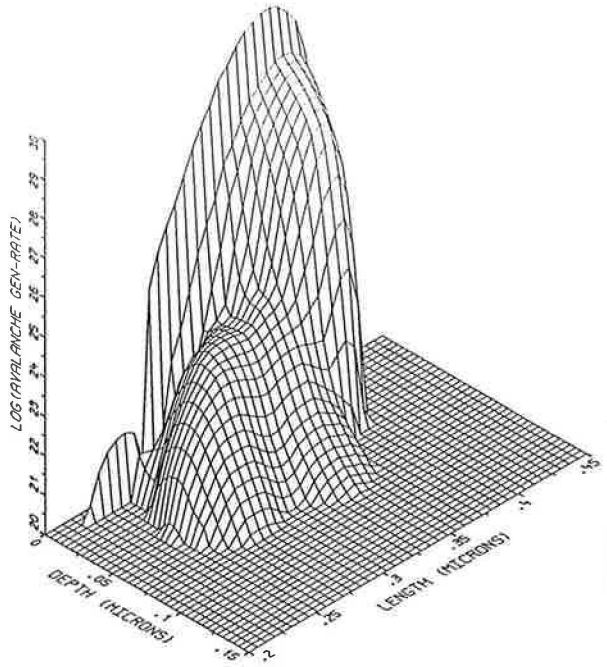

Fig.14: Impact Ionization Rate at 77K (UG=2V, UD=2V)

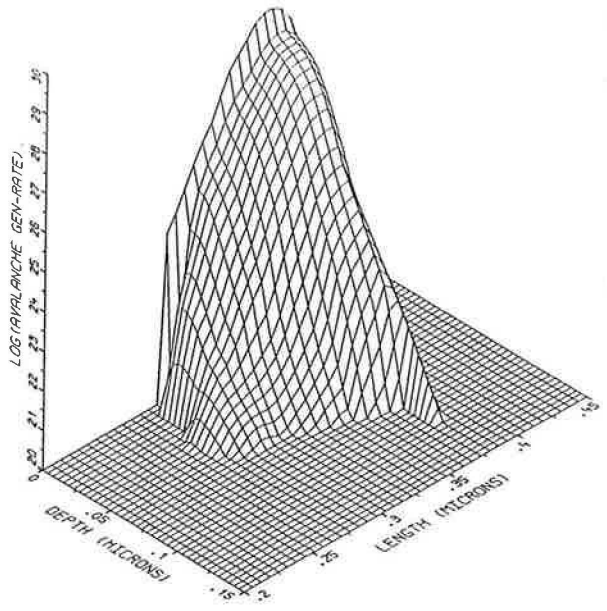

Fig.15: Impact Ionization Rate at 300K (UG=2V, UD=2V)

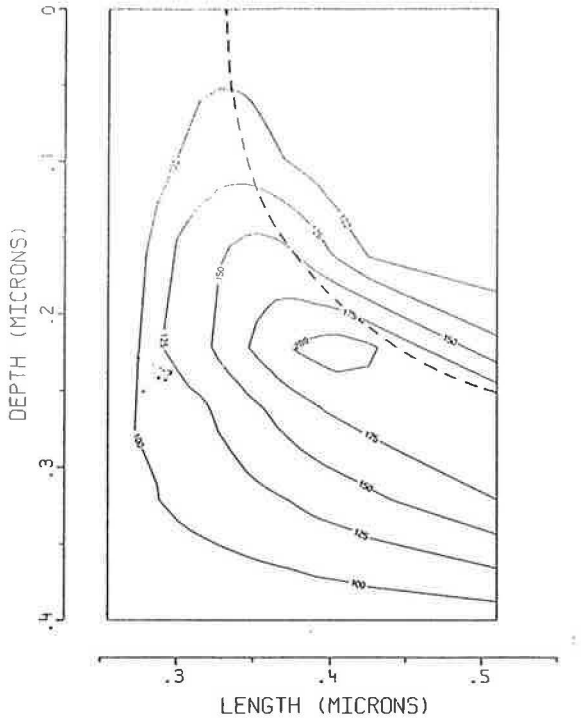

Fig.16: Electron Temperature at 77K (UG=2V, UD=2V)

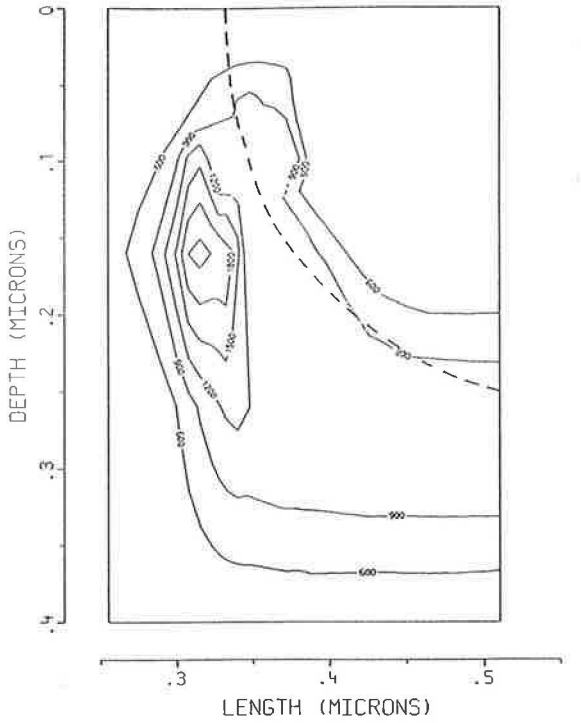

Fig.17: Electron Temperature at 300K (UG=2V, UD=2V)

Fig.16 and Fig.17 show the distribution of electron temperature at 77K and 300K ambient temperature, respectively. The maximum temperature is 213K at 77K and 2220K at room temperature. This maximum is in both cases located in the reverse biased drain substrate diode with a smooth transition into the channel. The position of the maximum is deeper in the substrate and closer to the drain area at 77K compared to the room temperature result. It is worthshile to note that the channel charge stays nearly thermal close to the interface. The phenomenon of smaller carrier heating at liquid nitrogen temperature is a result of a smoother distribution and a smaller maximum of the driving force. Smaller carrier heating at 77K has been confirmed by many simulations. However, remembering various private communications it is not really expected, particularly in view of larger energy relaxation times at liquid nitrogen temperature.

One numerical experiment carried out concerns scaling. The device under investigation has been scaled up by a factor of $3,9 = 300K/77K$ using the classical MOS device scaling rules. Fig.18 shows the simulated output characteristics for the up-scaled device We can see that such a temperature scaling is relatively crude. The current drive capability is not achieved, since the threshold voltage does not scale [63]. Furthermore, avalanche breakdown is more dominant indicated by the bending of the characteristics. The output conductance in the saturation region, however, scales very well.

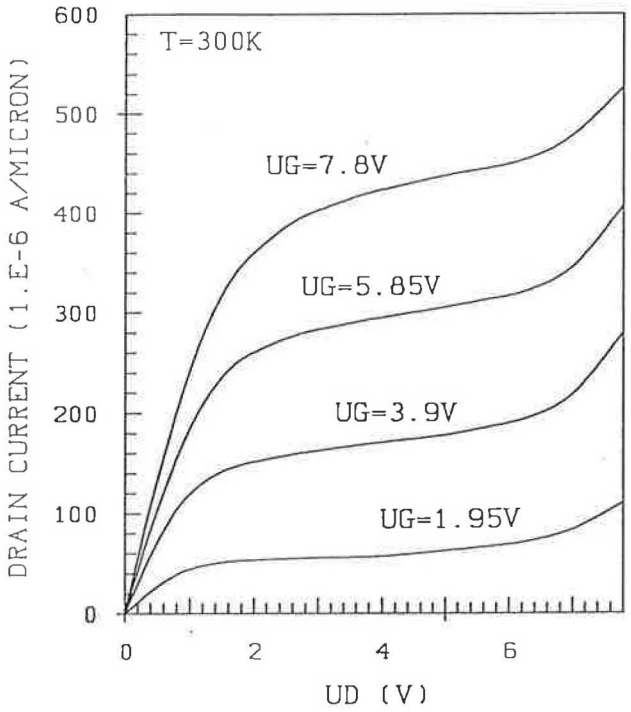

Fig.18: Simulated Output Characteristics at 300K for Up-Scaled Device

Similar investigations for a lightly doped drain technology can be found in [59].

The question remains how good these simulation results agree with measurements. The device presented has not been fabricated with 0,51 micron gate length. Satisfactory

agreement has been achieved for devices down to 3/4 micron gate length. To be able to judge rigorously the agreement between measurement and simulation at low temperatures one should also look at results obtained with different programs. These can be found, e.g., in [35] for a modified version of CADDET, in [52], [65] for a modified version of FIELDAY and in [69] for a modified version of GEMINI.

ACKNOWLEDGMENT

This work is considerably supported by the research laboratories of SIEMENS AG at Munich, FRG, the research laboratories of DIGITAL EQUIPMENT CORPORATION at Hudson, USA. I am indebted to Prof.H.Pötzl for many critical and stimulating discussions.

REFERENCES

[1] Ahmad N., Arora V.K.,
Velocity-Field Profile of n-Silicon: A Theoretical Analysis,
IEEE Trans.Electron Devices, Vol.ED-33, pp.1075-1077, 1986.

[2] Ali-Omar M., Reggiani L.,
Drift and Diffusion of Charge Carriers in Silicon and Their Empirical Relation to the Electric Field,
Solid-State Electronics, Vol.30, pp.693-697, 1987.

[3] Aoki M., Yano K., Masuhara T., Ikeda S., Meguro S.,
Optimum Crystallographic Orientation of Submicron CMOS Devices,
Proc.IEDM, pp.577-580, 1985.

[4] Aoki M., Hanamura S., Masuhara T., Yano K.,
Performance and Hot-Carrier Effects of Small CRYO-CMOS Devices,
IEEE Trans.Electron Devices, Vol.ED-34, pp.8-18, 1987.

[5] Aoki M., Yano K., Masuhara T., Ikeda S., Meguro S.,
Optimum Crystallographic Orientation of Submicrometer CMOS Devices Operated at Low Temperatures,
IEEE Trans.Electron Devices, Vol.ED-34, pp.52-57, 1987.

[6] Arora N.D., Hauser J.R., Roulston D.J.,
Electron and Hole Mobilities in Silicon as a Function of Concentration and Temperature,
IEEE Trans.Electron Devices, Vol.ED-29, pp.292-295, 1982.

[7] Arora N.D., Gildenblat G.S.,
A Semi-Emperical Model of the MOSFET Inversion Layer Mobility for Low-Temperature Operation,
IEEE Trans.Electron Devices, Vol.ED-34, pp.89-93, 1987.

[8] Baccarani G., Wordeman M.R.,
Transconductance Degradation in Thin-Oxide MOSFET's,
Proc.IEDM, pp.278-281, 1982.

[9] Baccarani G.,
Physics of Submicron Devices,
Proc.VLSI Process and Device Modeling, pp.1-23, Katholieke Universiteit Leuven, 1983.

[10] Baccarani G., Wordeman M.R.,
Transconductance Degradation in Thin-Oxide MOSFET's,
IEEE Trans.Electron Devices, Vol.ED-30, pp.1295-1304, 1983.

[11] Baccarani G., Wordeman M.R.,
An Investigation of Steady-State Velocity Overshoot in Silicon,
Solid-State Electronics, Vol.28, pp.407-416, 1985.

[12] Blakemore J.S.,
Approximations for Fermi-Dirac Integrals, especially the Function F1/2(x) used to Describe Electron Density in a Semiconductor,
Solid-State Electronics, Vol.25, pp.1067-1076, 1982.

[13] Blotekjaer K.,
Transport Equations for Electrons in Two-Valley Semiconductors,
IEEE Trans.Electron Devices, Vol.ED-17, pp.38-47, 1970.

[14] Canali C., Ottaviani G.,
Saturation Values of the Electron Drift Velocity in Silicon between 300K and 4.2K,
Physics Lett., Vol.32A, pp.147-148, 1970.

[15] Canali C., Majni G., Minder R., Ottaviani G.,
Electron and Hole Drift Velocity Measurements in Silicon and Their Empirical Relation to Electric Field and Temperature,
IEEE Trans.Electron Devices, Vol.ED-22, pp.1045-1047, 1975.

[16] Caughey D.M., Thomas R.E.,
Carrier Mobilities in Silicon Empirically Related to Doping and Field,
Proc.IEEE, Vol.52, pp.2192-2193, 1967.

[17] Chrzanowska-Jeske M., Jaeger R.C.,
Modeling of Temperature Dependent Transport Parameters for Low Temperature Bipolar Transistor Simulation,
Proc.Symposium on Low Temperature Electronics and High Temperature Superconductors, The Electrochemical Society, Vol.88-9, pp.30-38, 1988.

[18] Cody W.J., Thacher H.C.,
Rational Chebyshev Approximations for Fermi-Dirac Integrals of Orders -1/2, 1/2 and 3/2,
Math.Comp., Vol.21, pp.30-40, 1967.

[19] Colonna-Romano L.M., Deverell D.R.,
Operation of a CMOS Microprocessor While Immersed in Liquid Nitrogen,
IEEE J.Solid-State Circuits, Vol.SC-21, pp.491-492, 1986.

[20] Crowell C.R., Sze S.M.,
Temperature Dependence of Avalanche Multiplication in Semiconductors,
Appl.Phys.Lett., Vol.9, pp.242-244, 1966.

[21] DeMari A.,
An Accurate Numerical Steady-State One-Dimensional Solution of the P-N Junction,
Solid-State Electronics, Vol.11, pp.33-58, 1968.

[22] Debye P.P., Conwell E.M.,
Electrical Properties of N-Type Germanium,
Physical Review, Vol.93, pp.693-706, 1954.

[23] Decker D.R., Dunn C.N.,
Temperature Dependence of Carrier Ionization Rates and Saturated Velocities in Silicon,
J.Electronic Mat., Vol.4, pp.527-547, 1975.

[24] Dorkel J.M., Leturcq Ph.,
Carrier Mobilities in Silicon Semi-Empirically Related to Temperature, Doping and Injection Level,
Solid-State Electronics, Vol.24, pp.821-825, 1981.

[25] Duke D.W.,
Use of the ETA-10 Supercomputer: A Status Report,
Proc.Symposium on Low Temperature Electronics and High Temperature Superconductors, The Electrochemical Society, Vol.88-9, pp.30-38, 1988.

[26] Dziewior J., Schmid W.,
Auger Coefficients for Highly Doped and Highly Excited Silicon,
Appl.Phys.Lett., Vol.31, pp.346-348, 1977.

[27] Faricelli J.,
Private Communication,
1987.

[28] Frey J.,
Transport Physics for VLSI,
in: Introduction to the Numerical Analysis of Semiconductor Devices and Integrated Circuits, pp.51-57, Boole Press, Dublin 1981.

[29] Gaensslen F.H., Jaeger R.C., Walker J.J.,
Low-Temperature Threshold Behavior of Depletion Mode Devices,
Proc.IEDM, pp.520-524, 1976.

[30] Gaensslen F.H., Rideout V.L., Walker E.J., Walker J.J.,
Very Small MOSFET's for Low Temperature Operation,
IEEE Trans.Electron Devices, Vol.ED-24, pp.218-229, 1977.

[31] Gaensslen F.H., Jaeger R.C.,
Temperature Dependent Threshold Behaviour of Depletion Mode MOSFET's,
Solid-State Electronics, Vol.22, pp.423-430, 1979.

[32] Gaensslen F.H., Jaeger R.C.,
Behavior of Electrically Small Depletion Mode MOSFET's at Low Temperature,
Solid-State Electronics, Vol.24, pp.215-220, 1981.

[33] Gummel H.K.,
A Self-Consistent Iterative Scheme for One-Dimensional Steady State Transistor Calculations,
IEEE Trans.Electron Devices, Vol.ED-11, pp.455-465, 1964.

[34] Hänsch W., Selberherr S.,
MINIMOS 3: A MOSFET Simulator that Includes Energy Balance,
IEEE Trans.Electron Devices, Vol.ED-34, pp.1074-1078, 1987.

[35] Henning A.K., Chan N., Plummer J.D.,
Substrate Current in n-Channel and p-Channel MOSFET's between 77K and 300K: Characterization and Simulation,
Proc.IEDM, pp.573-576, 1985.

[36] Henning A.K., Chan N.N., Watt J.T., Plummer J.D.,
Substrate Current at Cryogenic Temperatures: Measurements and a Two-Dimensional Model for CMOS Technology,
IEEE Trans.Electron Devices, Vol.ED-34, pp.64-74, 1987.

[37] Hess K., Iafrate G.J.,
Theory and Applications of Near Ballistic Transport in Semiconductors,
Proc.IEEE, Vol.76, pp.519-532, 1988.

[38] Heywang W., Pötzl H.,
Bandstruktur und Stromtransport,
Springer, Berlin, 1976.

[39] Hiroki A., Odanaka S., Ohe K., Esaki H.,
A Mobility Model for Submicrometer MOSFET Device Simulations,
IEEE Electron Device Lett., Vol.EDL-8, pp.231-233, 1987.

[40] Jaeger R.C.,
Private Communication,
October 1987.

[41] Jaeger R.C., Gaensslen F.H.,
Low Temperature MOS Microelectronics,
Proc.Symposium on Low Temperature Electronics and High Temperature Superconductors, The Electrochemical Society, Vol.88-9, pp.43-54, 1988.

[42] Jaggi R., Weibel H.,
High-Field Electron Drift Velocities and Current Densities in Silicon,
Helv.Phys.Acta, Vol.42, pp.631-632, 1969.

[43] Jaggi R.,
High-Field Drift Velocities in Silicon and Germanium,
Helv.Phys.Acta, Vol.42, pp.941-943, 1969.

[44] Kamgar A.,
Miniaturization of Si MOSFET's at 77K,
IEEE Trans.Electron Devices, Vol.ED-29, pp.1226-1228, 1982.

[45] Kinugawa M., Kakumu M., Usami T., Matsunaga J.,
Effects of Silicon Surface Orientation on Submicron CMOS Devices,
Proc.IEDM, pp.581-584, 1985.

[46] Lau D., Gildenblat G., Sodini G.G., Nelson D.E.,
Low-Temperature Substrate Current Characterization of n-Channel MOSFET's,
Proc.IEDM, pp.565-568, 1985.

[47] Li S.S., Thurber W.R.,
The Dopant Density and Temperature Dependence of Electron Mobility and Resistivity in n-Type Silicon,
Solid-State Electronics, Vol.20, pp.609-616, 1977.

[48] Markowich P.A.,
The Stationary Semiconductor Device Equations,
Springer, Wien New-York, 1986.

[49] Nishida T., Sah C.T.,
A Physically Based Mobility Model for MOSFET Numerical Simulation,
IEEE Trans.Electron Devices, Vol.ED-34, pp.310-320, 1987.

[50] Okuto Y., Crowell C.R.,
Threshold Energy Effect on Avalanche Breakdown Voltage in Semiconductor Junctions,
Solid-State Electronics, Vol.18, pp.161-168, 1975.

[51] Robertson P.J., Dumin D.J.,
Ballistic Transport and Properties of Submicrometer Silicon MOSFET's from 300 to 4.2K,
IEEE Trans.Electron Devices, Vol.ED-33, pp.494-498, 1986.

[52] Sai-Halasz G.A.,
Processing and Characterization of Ultra Small Silicon Devices,
Proc.ESSDERC Conf., pp.71-80, 1987.

[53] Schütz A., Selberherr S., Pötzl H.,
Analysis of Breakdown Phenomena in MOSFET's,
IEEE Trans.CAD of Integrated Circuits and Systems, Vol.CAD-1, pp.77-85, 1982.

[54] Seavey M.,
Private Communication,
1987.

[55] Seeger K.,
Semiconductor Physics,
Springer, Wien, 1973.

[56] Selberherr S., Schütz A., Pötzl H.,
MINIMOS - A Two-Dimensional MOS Transistor Analyzer,
IEEE Trans.Electron Devices, Vol.ED-27, pp.1540-1550, 1980.

[57] Selberherr S.,
Analysis and Simulation of Semiconductor Devices,
Springer, Wien New-York, 1984.

[58] Selberherr S., Griebel W., Pötzl H.,
Transport Physics for Modeling Semiconductor Devices,
in: Simulation of Semiconductor Devices and Processes, pp.133-152, Pineridge Press, Swansea, 1984.

[59] Selberherr S.,
Low Temperature Mos Device Modeling,
Proc.Symposium on Low Temperature Electronics and High Temperature Superconductors, The Electrochemical Society, Vol.88-9, pp.43-86, 1988.

[60] Selberherr S.,
MOS Device Modeling at 77K,
IEEE Trans.Electron Devices, Vol.ED-36, pp.1464-1474, 1989.

[61] Selberherr S., Hänsch W., Seavey M., Slotboom J.,
The Evolution of the MINIMOS Mobility Model,
Solid-State Electronics, Vol.33, pp.1425-1436, 1990.

[62] Shahidi G.G., Antoniadis D.A., Smith H.I.,
Electron Velocity Overshoot at 300K and 77K in Silicon MOSFET's with Submicron Channel Length,
Proc.IEDM, pp.824-825, 1986.

[63] Solomon P.M.,
Options for High Speed Logic at 77K,
Proc.Symposium on Low Temperature Electronics and High Temperature Superconductors, The Electrochemical Society, Vol.88-9, pp.3-17, 1988.

[64] Sugano T.,
Low Temperature Electronics Research in Japan,
Proc.Symposium on Low Temperature Electronics and High Temperature Superconductors, The Electrochemical Society, Vol.88-9, pp.18-29, 1988.

[65] Sun Y-C.J., Taur Y., Dennard R.H., Klepner S.P.,
Submicrometer-Channel CMOS for Low-Temperature Operation,
IEEE Trans.Electron Devices, Vol.ED-34, pp.19-27, 1987.

[66] Sutherland A.D.,
An Improved Empirical Fit to Baraff's Universal Curves for the Ionization Coefficients of Electron and Hole Multiplication in Semiconductors,
IEEE Trans.Electron Devices, Vol.ED-27, pp.1299-1300, 1980.

[67] Sze S.M.,
Physics of Semiconductor Devices,
Wiley, New York, 1969.

[68] Watt J.T., Fishbein B.J., Plummer J.D.,
A Low-Temperature NMOS Technology with Cesium-Implanted Load Devices,
IEEE Trans.Electron Devices, Vol.ED-34, pp.28-38, 1987.

[69] Woo J.C.S., Plummer J.D.,
Short Channel Effects in MOSFET's at Liquid-Nitrogen Temperature,
IEEE Trans.Electron Devices, Vol.ED-33, pp.1012-1019, 1986.