JOURNAL OF OPTOELECTRONICS AND ADVANCED MATERIALS Vol. 9, No. 9, September 2007, p. 2703 - 2708

Threshold voltage in MOSFETs and MODFETs as a problem of nonlinear dynamics

RAJKO M. ŠAŠIĆ, PETAR M. LUKIĆ[a], RIFAT M. RAMOVIĆ[b], STANKO M. OSTOJIĆ
Faculty of Technology and Metallurgy, University of Belgrade, Serbia
[a]Faculty of Mechanical Engineering, University of Belgrade, Serbia
[b]Faculty of Electrical Engineering, University of Belgrade, Serbia

A widely known (and also very well examined) problem of threshold voltage in MOSFETs and MODFETs has been described with the tools of nonlinear dynamics. Some interesting features of such devices have been derived, together with the outstanding differences between them.

(Received May 25, 2007; accepted August 24, 2007)

Keywords: Phase portrait, Nonlinear dynamics, Fixed points, Stability, MOSFET, MODFET, Threshold voltage

1. Introduction

The conventional Si MOSFETs, as well as GaAs (and other heterostructure friendly materials also) MODFETs, are very well known to have a following property: in both types of devices an extrapolated threshold voltage can be recognized and calculated [1]. Even more, it appears as a qrucial phusical quantity that governs their operation [3]. Shortly speaking, in such structures a quantum well arises, in which electrons from other regions of a device gather [2, 3]. The electrons surface concentration is controlled by an external (lateral) gate voltage V_g. In the case of conventional Si MOSFETs this influence can be expressed through following expressions [7]:

$$V_g = V_{FB} + \frac{\sqrt{2e\varepsilon N_A}}{d_{ox}'}\sqrt{\psi_s + \phi_t e^{(\psi_s - 2\phi_F)/\phi_t}} + \psi_s \quad (1)$$

$$n_s = \sqrt{2e\varepsilon N_A}\left(\sqrt{\psi_s + \phi_t e^{(\psi_s - 2\phi_F)/\phi_t}} - \sqrt{\psi_s}\right) \quad (2)$$

where ψ_s denotes electrostatic potential at the top of the channel, V_{FB} denotes flat-band voltage, ε is the oxide dielectric permeability and d_{ox}' its thickness, ϕ_t=kT/*e*, ϕ_F=E_F/*e* and *e* is the electron charge. Obviously, for low values of V_g, the exponential term in both previous equations becomes negligible and the surface carriers concentration n_s in the channel remains close to zero. But for higher values of V_g, the exponential term becomes dominant (its dramatical nonlinearity plays the crucial role) and the dependence $n_s(V_g)$ becomes linear [3, 6]:

$$n_s = C_{ox}'\left(V_g - V_{T0}\right), \quad C_{ox}' = \frac{\varepsilon}{d_{ox}'} \quad (3)$$

The quantity V_{T0} is called the extrapolated threshold voltage and its extraction from equations (1) and (2) can be discussed in details. The point of this paper is the fact that this threshold voltage exists and therefore its detailled discussion will be omitted here. For the purpose of this paper it is quite sufficient to illustrate the dependence $n_s(V_g)$ as done in Fig. 1. Although the precise plot is smooth (it has a continuous first derivative), widely exploited characteristic is piecewise, i.e. considered to consist of two straight lines (sections) which independently describe regions of weak and strong inversion. It is also considered that nothing dramatic happens when V_g is further increased (naturally in the usually exploited range) [3, 7].

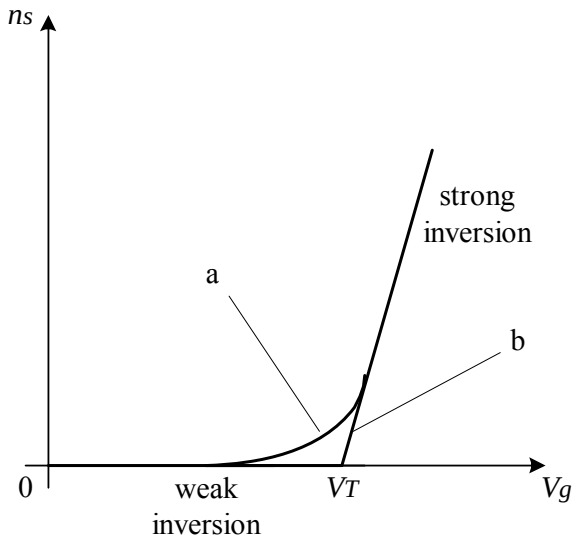

Fig. 1. Surface concentration of carriers in the channel n_s versus gate voltage V_g: a - exact solution; b – piecewise approximation for Si MOSFETs

Let us now turn to heterostructure devices and investigate AlGaAs/GaAs heterojunction as its representative sample. The equations responsible for the evaluation of carriers' concentration in the channel are [5]:

$$n_s = \frac{m^* kT}{\pi\hbar^2} \ln\left[\left(1+e^{\frac{E_{Fi}-E_1}{kT}}\right)\left(1+e^{\frac{E_{Fi}-E_2}{kT}}\right)\right] \quad (4)$$

$$E_1 = \gamma_1 n_s^{2/3} \quad (4a)$$

$$E_2 = \gamma_2 n_s^{2/3} \quad (4b)$$

$$n_s = \frac{\varepsilon_2}{d_2}\left(V_P - V_B + V_g + \frac{\Delta E_c}{e} - \frac{E_{Fi}(n_s)}{e}\right) \quad (5)$$

In equation (4) E_1, and E_2 are the lowest two quantum levels (the higher ones are supposed not to play any role), while γ_1 and γ_2 are usually adjusted according to Schubnikov de Haas experimental data. Equation (5) is derived assuming that AlGaAs layer is entirely depleted. To support these explanations we usually give a glance to Fig. 2.

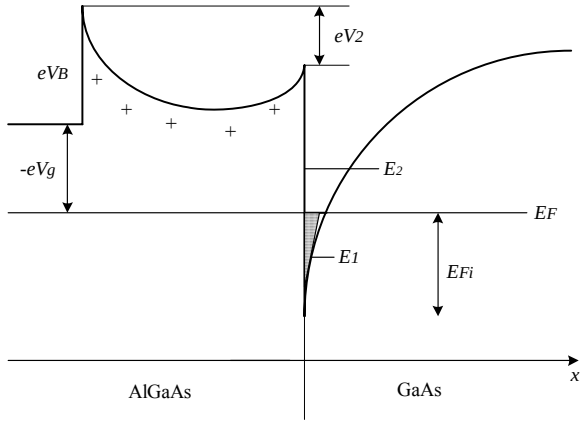

Fig. 2. Band diagram of AlGaAs/GaAs heterojunction with an applied external voltage V_g.

The equation (5) is a transcedental one, since E_{Fi} is a function of n_s. If we ignore its variation, the following linear relations are obtained [5, 6]:

$$n_s = 0; \quad V_g < V_T \quad (6a)$$

$$n_s = \frac{\varepsilon_2}{ed_2}(V_g - V_T); \quad V_T < V_g < V_m \quad (6b)$$

$$V_T = V_B - V_F - \frac{\Delta E_c}{e} + \frac{E_F(n_{sc})}{e} \quad (6c)$$

The quantity V_T is again recognized as an extrapolated thresholdvoltage, similar to conventional Si MOSFETs. The increase of positive V_g decreases the number of uncompensated donors imaging the charge on the metal and therefore increases the fraction of donors contributing electrons to the two-dimensional electron gas in the channel. But the further increase of V_g brings the departure of heterostructure devices' behavior from Si MOSFET's ones. For V_g values larger than $(en_{so}d_2/\varepsilon_2)=V_m$ the depletion regions near metal (Schottky junction) and two-dimensional electron gas (channel) cease to overlap (they become separate). While the depletion approximation is no longer valid across the entire AlGaAs layer, the equation (5) becomes unconvenient. The control of n_s through gate voltage V_g is lost and surface concentration of carriers saturates toward its equilibrium value n_{so} [2]:

$$n_{so} = \sqrt{2\varepsilon_2 N_D[\Delta E_c - \delta - E_{Fi}(n_{so})] + (eN_D)^2} - eN_D \quad (7)$$

where δ is the quantity well known from literature. Finally, the dependence $n_s(V_g)$ is given in Fig. 3 [1].

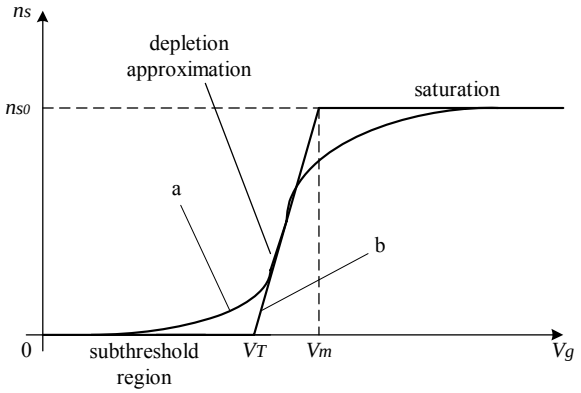

Fig. 3. Surface concentration of carriers n_s versus applied gate voltage V_g: a - exact solution; b - piecewise approximation in AlGaAs/GaAs heterojunction.

The reason why the plots in Figs. 1 and 3 differ for higher values of V_g can be found in the fact that in conventional Si MOSFETs electrons come only from buffer layer, while in heterostructure devices they can be supplied either from buffer layer or from AlGaAs layer. The Fig. 1 and 3 will be the starting point for our further analysis.

2. Dynamical picture

The question arised if it was possible to investigate the role that Fig. 1 and 3 could play in the dynamical treatment of carriers' surface concentration evolution. Therefore it is assumed that gate voltage is switched on in

an arbitrary moment, i.e. its time dependence is expressed by step function:

$$V_g(t) = V_g h(t) \tag{8}$$

Then we shall make an attempt to construct the (nonlinear) first order differential equation that describes carriers' surface concentration evolution (together with initial condition):

$$\frac{dn_s}{dt} = \dot{n}_s = f(n_s); \quad t \geq 0 \tag{9a}$$

$$n_s(0) = K > 0 \tag{9b}$$

The next step is to connect equation (9a) with the Figs. 1 and 3 in the following manner. The solutions of the equation:

$$f(n_s) = 0 \tag{10}$$

are simply plotted in these figures. Gate voltage appears as a parameter whose magnitude defines the belonging to different sections. The relation (10) simultaneously demands $\dot{n}_s$ =0. This fact was a hint to assume that the solutions of the equation (10) could actually play the role of asymptotic solutions of the equation (9a) when time tends to infinity [4].

In the case of conventional Si MOSFETs, we try with square dependence $f(n_s)$:

$$\dot{n}_s = \lambda^2(V_g - V_T)n_s - \beta^2 {n_s}^2 = f(n_s) \tag{11}$$

where λ^2, β^2 denote positive constants and V_T is the extrapolated threshold voltage described in the previous section. The integration of equation (11) is straightforward and, together with the initial condition $n_s(0)$=K, gives final result:

$$\ln\left\{\left|\frac{n_s}{K}\right| \cdot \frac{\left|\frac{\lambda^2(V_g - V_T)}{\beta^2} - K\right|}{\left|\frac{\lambda^2(V_g - V_T)}{\beta^2} - n_s\right|}\right\} = \lambda^2(V_g - V_T)\cdot t \tag{12a}$$

$$\text{i.e.} \quad \frac{|n_s|}{|K|} \cdot \frac{\left|\lambda^2(V_g - V_T) - \beta^2 K\right|}{\left|\lambda^2(V_g - V_T) - \beta^2 n_s\right|} = e^{\lambda^2(V_g - V_T)t} \tag{12b}$$

which is valid for each positive V_g except V_g=V_T. The straightforward conclusion follows:

- V_g<V_T $\Rightarrow$

$$\lim_{t\to+\infty} \frac{|n_s|}{|K|} \cdot \frac{\left|\lambda^2(V_g - V_T) - \beta^2 K\right|}{\left|\lambda^2(V_g - V_T) - \beta^2 n_s\right|} = 0 \tag{13a}$$

and therefore should be:

$$\lim_{t\to+\infty} n_s(t) = 0 \tag{13b}$$

- V_g>V_T

$$\Rightarrow \lim_{t\to+\infty} \frac{|n_s|}{|K|} \cdot \frac{\left|\lambda^2(V_g - V_T) - \beta^2 K\right|}{\left|\lambda^2(V_g - V_T) - \beta^2 n_s\right|} \to +\infty \tag{14a}$$

which implies:

$$\lim_{t\to+\infty} \left(\lambda^2(V_g - V_T) - \beta^2 n_s\right) = 0 \tag{14b}$$

$$\lim_{t\to+\infty} n_s(t) = \frac{\lambda^2}{\beta^2}(V_g - V_T) \tag{14c}$$

Obviously, it has been demonstrated that asymptotic solutions of the equation (11) were invariant under the variation of initial condition. These asymptotic solutions, given in expressions (13) and (14), are exactly two linear sections of the piecewiese characteristic plotted in Fig. 1 [4].

Let us now see what can be claimed about this problem from the point of view of nonlinear dynamics. First of all, we have to sketch a phase portrait (Fig. 4).

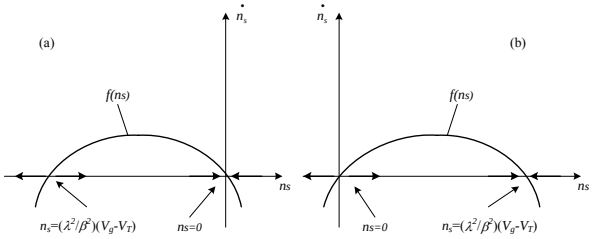

Fig. 4. Phase portrait $\dot{n}_s$ =f(n_s) in two different cases: (a) (V_g-V_T)<0 and (b) (V_g-V_T)>0 for Si MOSFETs

The solutions of equation (10) are called fixed points (as signed in the Fig. 4). Without solving differential equation (11), one can immediately conclude that if (V_g-V_T)<0 (figure 4a) n_s=(λ^2/β^2)(V_g-V_T) is an unstable fixed point, while n_s=0 is a stable fixed point (observe the direction of the arrows in Fig. 4. – into the point/out of the point). But the increase of V_g over V_T brings a dramatic change: a fixed point n_s=0 loses its stability and a new stable fixed point appears at n_s=(λ^2/β^2)(V_g-V_T). In both cases a stable fixed point should be recognized as an

asymptotic solution that does not depend on initial condition. It is also obvious that (λ^2/β^2) will be the slope of the right section ($V_g>V_T$) in the Fig. 1, while α^2 controls how fast is this asymptotic solution achieved. Last but not least, the only physical meaning have the values $n_s>0$.

The question of AlGaAs/GaAs heterostructures (as well as other similar structures) is more complex due to the Fig. 3. In order to handle this characteristic, one must use cubic polynomial function:

$$\dot{n}_s = \alpha^2 n_s\left(n_s - n_{so}x\right)\left(n_s - n_{so}\right), \quad n_s(0) = K \tag{15}$$

with: $x=(V_g-V_T)/(V_m-V_T)$, while V_m and n_{so} are defined according to equation (7) and Fig. 3. The integration of equation (15) can be carried out analytically:

$$\frac{1}{x}\ln\left|\frac{n_s}{K}\right| + \frac{1}{x(x-1)}\ln\left|\frac{n_s - n_{so}x}{K - n_{so}x}\right| - \frac{1}{x-1}\ln\left|\frac{n_s - n_{so}}{K - n_{so}}\right| = n_{so}^2\alpha^2 t \tag{16a}$$

or:
$$\left|\frac{n_s}{K}\right|^{\frac{1}{x}} \cdot \left|\frac{n_s - n_{so}x}{K - n_{so}x}\right|^{\frac{1}{x(x-1)}} \cdot \left|\frac{n_s - n_{so}}{K - n_{so}}\right|^{\frac{1}{1-x}} = e^{n_{so}^2\alpha^2 t} \tag{16b}$$

Depending on parameter x different possibilities arise:

- $x<0$ ($V_g<V_T$), $\lim_{t\to+\infty} e^{n_{so}^2\alpha^2 t} \to +\infty$

because $\frac{1}{x} < 0$, $\frac{1}{x(x-1)} > 0$, $\frac{1}{1-x} > 0$

the only possibility is:

$$\lim_{t\to+\infty} n_s(t) = 0 \tag{17}$$

- $0<x<1$ ($V_T<V_g<V_m$), again $\lim_{t\to+\infty} e^{n_{so}^2\alpha^2 t} \to +\infty$

because $\frac{1}{x} > 0$, $\frac{1}{x(x-1)} < 0$, $\frac{1}{1-x} > 0$

the possibility that survives is:

$$\lim_{t\to+\infty} n_s(t) = n_{s0}x \tag{18}$$

- $x>1$ ($V_g>V_m$), and $\lim_{t\to+\infty} e^{n_{so}^2\alpha^2 t} \to +\infty$ once more:

implies $\frac{1}{x} > 0$, $\frac{1}{x(x-1)} > 0$, $\frac{1}{1-x} < 0$

therefore according to equation (16b) must be valid:

$$\lim_{t\to+\infty} n_s(t) = n_{so} \tag{19}$$

It has been again demonstrated that asymptotic solutions of the equation (15) did not depend on initial condition, but they appeared to coincide with three linear sections of the piecewise characteristic plotted in Fig. 3.

Now, we will concentrate ourselves on phase portraits in this case.

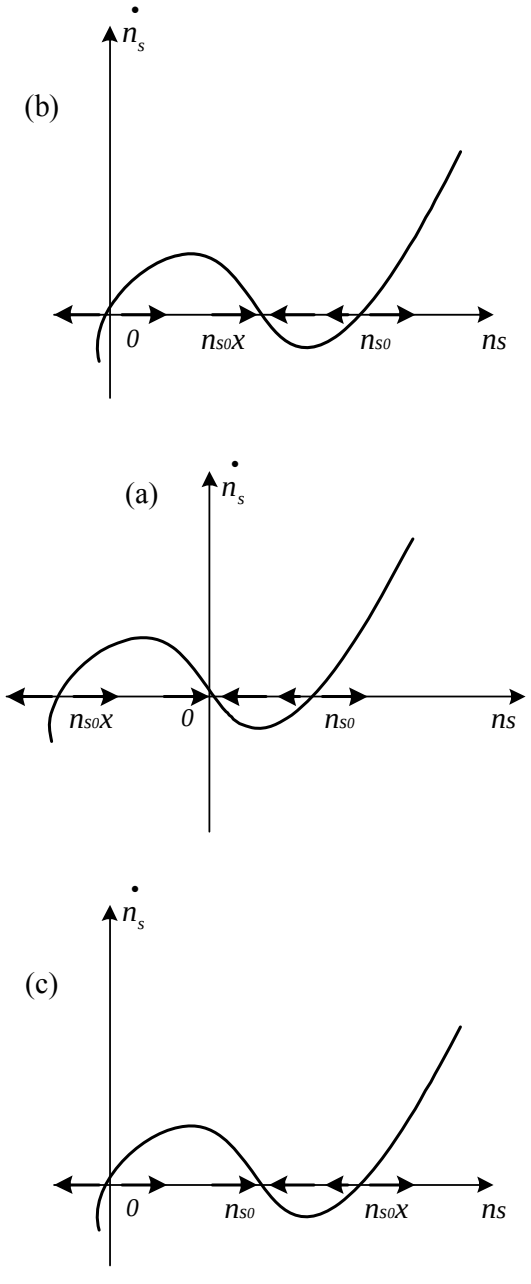

Fig. 5. *Phase portrait $\dot{n}_s = f(n_s)$ in three different cases: (a) $V_g<V_T$ (b) $V_T<V_g<V_m$ (c) $V_g>V_m$ for AlGaAs/GaAs heterostructure MODFETs.*

According to equation (15) in this case three fixed points exist. From Fig. (5a) one can make a conclusion that for (V_g<V_T) n_s=0 is the only stable fixed point, while n_s=$n_{so}x$ and n_s=n_{so} are unstable. Fig. (5b) implies that in the region (V_T<Vg<V_m) the stable fixed point is $n_{so}x$ and the other two are unstable. It is worth mentioning that in this section the asymptotic solution of equation (15) is an increasing linear function of V_g. For even larger values of gate voltage (V_g>V_m) the previously described saturation appears and therefore the stable fixed point is posed at n_{so}. Other two fixed points (n_s=0 and $n_{so}x$) are now unstable.

Similar to the case of conventional Si MOSFETs, for MODFETs the asymptotic solution tends to the stable fixed point over the whole range of applied voltages. Important to say, this can be concluded analysing the structure of equation (15) in terms of nonlinear dynamics without its solving.

In the approach described above, the values V_g=V_T for Si MOSFETs, as well as V_g=V_T or V_g=V_m for heterostructure MODFETs, can appear as "critical" (or "singular") points. Firstly, it would be very easy to show that for these values of gate voltage the asymptotic solutions tend to expected values that belong to the $n_s(V_g)$ characteristic from Figs. (1) and (3). Secondly, but more important, these piecewise characteristics are the approximations to the more accurate continuous ones, so such treatment at values V_g=V_T or V_g=V_m loses sense. At the end of this section, it is necessary to mention that only the region n_s>0 has physical meaning [6].

3. Discussion

To encircle the whole story it would be very useful if the parameters exploited in our description were given some physical explanation. In order to achieve that in the case of Si MOSFETs let's rewrite equation (11) in the following form:

$$\dot{n}_s = \left(\lambda^2 V_g - \beta^2 n_s\right) n_s - \lambda^2 V_T n_s \tag{20}$$

The second term on the right side of equation (20) can help us to recognize the inverse lifetime of carriers in the channel:

$$\lambda^2 V_T = \frac{1}{\tau} \tag{21}$$

This lifetime is also the mean value of channel duration, i.e. it emphasizes the fact that channel will disappear as time goes by if no external gate voltage is applied. The analysis of the first term in equation (20) suggests that this gate voltage V_g>0 is necessary for channel creation, because it causes the quantum well at the interface to be formed. The increase of V_g makes this quantum well deeper. But if carriers gather in this well they efficiently diminish the influence of V_g increase and that is the explanation for the appearance of the entire first term in equation (20). Now equation (20) can be rewritten in the final form:

$$\dot{n}_s = -\frac{1}{\tau} n_s + \left(\frac{V_g}{V_T}\frac{1}{\tau} - \beta^2 n_s\right) n_s \tag{21}$$

Parameter β^2 is simply a measure of the influence described above and at the moment it is left unexplained in details.

Equation (15), which is valid for heterostructure MODFETs, can be treated in a similar way, what after rearrangement gives:

$$\dot{n}_s = \alpha^2\left[n_{so}{}^2\frac{V_g}{V_m - V_T} - n_{so}\left(1 + \frac{V_g - V_T}{V_m - V_T}\right)n_s + n_s{}^2\right] n_s - \alpha^2 n_{so}{}^2 \frac{V_T}{V_m - V_T} n_s \tag{22}$$

The coefficient:

$$\alpha^2 n_{so}{}^2 \frac{V_T}{V_m - V_T} = \frac{1}{\tau} \tag{23}$$

will be again recognized as an inverse lifetime of carriers in channel with the same explanation as above. The first term in equation (22) is more complex than in the case of Si MOSFETs. If MODFETs are considered, the positive influence of V_g on channel creation will be suppressed by the increase of n_s for its lower values, but also encouraged by its even higher values because of $n_s{}^2$ (i.e. aggregate $n_s{}^3$) term. This complexity is a consequence of the fact that the carriers into MODFETs' channel can be brought either from AlGaAs layer or from GaAs bulk as mentioned before. The combination of equations (22) and (23) will give the final form:

$$\dot{n}_s = -\frac{1}{\tau} n_s + \left\{\frac{V_g}{V_T}\frac{1}{\tau} - \left[\alpha^2 n_{so} + \frac{1}{\pi n_{so}}\left(\frac{V_g}{V_T} - 1\right)\right] n_s + \alpha^2 n_s{}^2\right\} n_s \tag{24}$$

Sometimes it could be useful to introduce a normalized surface concentration of carriers in the channel $\eta_s(t)$:

$$n_s(t) = n_{so} \cdot \eta_s(t) \tag{25}$$

and its dynamics is described by following equation:

$$\dot{\eta}_s = \frac{1}{\tau}\left(\frac{V_g}{V_T} - 1\right)\eta_s - \frac{1}{\tau}\left[\left(\frac{V_m}{V_T} - 1\right) + \left(\frac{V_g}{V_T} - 1\right)\right]\eta_s{}^2 + \frac{1}{\tau}\left(\frac{V_m}{V_T} - 1\right)\eta_s{}^3 \tag{26}$$

This form is interesting because the coefficient α^2 appears explicitly in it no more. All parameters in it (except lifetime τ) are easily extractable from figure 3 (V_m, V_T, n_{so}).

4. Conclusions

Starting from the well known $n_s(V_g)$ characteristics in the case of either Si MOSFETs or heterostructure MODFETs corresponding first order differential equations have been constructed. Such an equation describes time evolution of carriers surface concentration in the channel, i.e. the evolution of the channel itself. Although the construction has been performed in the case of step function $V_g(t)$, the equation itself was believed to be valid for other types of $V_g(t)$ dependence too, what would significantly expand the field of its being useful to other nonstationary (or even periodic) regimes. This would perhaps supply a lot of relevant informations about the analysed structure.

Another thing to be emphasized is an outstanding nonlinearity met in the corresponding equations. This nonlinearity is so strong that the hypothetical linearization makes no sense. Therefore, it is very important to extract as much as possible informations by general methods for treatment of nonlinearities without direct solving of these equations.

At the end of the paper it can be concluded that as a key parameter for both types of devices appears lifetime of carriers' surface concentration in the channel. Possibly, if we knew exact (or even approximate) channel degradation mechanisms, we would be able to calculate this lifetime more or less successfully. And if we know the lifetime, we would be able to extract all missing data (parameters α^2, λ^2, β^2 from τ, as well as V_T, V_m, n_{so}) from the experimental data described by Figs. 1 and 3.

References

[1] Patric Roblin and Hans Rohdin: "High-speed heterostructure devices", Cambridge University Press, U. K., (2002).

[2] D. Delagebeaudeuf, N. Linh: "Metal (n) AlGaAs-GaAs two dimensional electron gas FET", IEEE Transaction on Electron Devices **ED-29**(6), 955-960 (1982).

[3] Yannis P. Tsividis: "Operation and Modeling of the MOS Transistors", McGraw Hill, (1988).

[4] Steven H. Strogatz: "Nonlinear Dynamics and Chaos (with applications to Physics, Biology, Chemistry and Engineering)", Westview Press, USA, (1994).

[5] P. M. Lukić, R. M. Ramović, R. M. Šašić: "Analytical Model of Electric Field in Heterojunction Region of HFET Structure", Journal of Optoelectronics and Advanced Materials **7**(3), 1611-1618 (2005).

[6] R. M. Šašić, P. M. Lukić, R. M. Ramović: "New Analytical HFET Current-Voltage (I-V) Characteristics Model", Journal of Optoelectronics and Advanced Materials **8**(1), 324-328 (2006).

[7] R. Šašić, R. M. Ramović, V. Herrmann: "A New Approximate Model for Short Channel MOS Devices", Phys. Stat. Sol. **165a**, 445-454 (1998).

*Corresponding author: plukic@mas.bg.ac.yu