https://doi.org/10.15407/ujpe66.7.625

M.V. STRIKHA,[1] A.I. KURCHAK [2]

[1] Taras Shevchenko National University of Kyiv,
Faculty of Radiophysics, Electronics, and Computer Systems
(4g, Akademika Glushkova Ave., Kyiv 03022, Ukraine; e-mail: maksym.strikha@gmail.com)
[2] V.E. Lashkaryov Institute of Semiconductor Physics, Nat. Acad. of Sci. of Ukraine
(41, Nauky Ave., Kyiv 03028, Ukraine)

FUNDAMENTAL LIMITS FOR THE MOSFET CONDUCTION CHANNEL LENGTH TAKING THE REAL PROFILE OF THE BARRIER POTENTIAL INTO ACCOUNT

The minimal length of the channel in the MOSFET, which is the principal device of modern electronics, has been estimated. The account of the real potential behavior in the channel demonstrates that, when some voltage is applied to the drain, electrons tunnel through a region that is essentially shorter than the physical channel length L. Therefore, the estimation of the minimal channel length in the Si-based MOSFET, which is available in the literature ($L_{\min} \approx 1.2$ nm), turns out substantially lowered. This discrepancy explains why, after having reached a working channel length of 5 nm, the value of 3 nm, which had been announced long ago, had not been achieved yet providing a proper level of the transistor functionality. The estimations made in this work confirm that the fundamental limits on the Si-based MOSFET scaling are currently almost reached.

K e y w o r d s: metal-oxide-semiconductor field-effect transistor, minimum channel length, barrier tunneling.

1. Introduction

Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) is a main device in the modern electronics. Therefore, it still remains the object of numerous experimental and theoretical studies (see, e.g., work [1]). The physics of the MOSFET is determined by the electron motion from the source S along the conduction channel to the drain D. The electron current I_D through the conduction channel of the transistor with the proper functionality level is effectively controlled by the gate G, which is insulated from the conduction channel by means of a dielectric layer. This control is based on the fact that the gate, depending on the applied voltage, changes the height of the potential barrier between the source and the drain.

A permanent miniaturization (scaling) of MOSFETs takes place. At the end of the 20th century, the standard length of the channel between the source and the drain was about 100 nm as a rule. Nowadays, it is about 10 nm. At the beginning of the 21st century, the creation of experimental specimens with much shorter channels was announced: first of an order of 6 nm [2], and afterward of 3 nm [3]. The thickness of the oxide layer in modern devices can already be narrower than 2 nm.

However, during the last two decades, it became more and more clear that the empirical "Moore's law", which satisfactorily described the situation with the transistor scaling since 1965 (the doubling of the number of elements on the chip every 24 months), is close to the exhaustion. This is not only due to numerous technological problems of scaling (see, e.g., work [4] and the references therein), but also to the fundamental limitations associated with the quantum-mechanical behavior of electrons in nanosystems.

2. Main Part

As long ago as in 1961, Rolf Landauer theoretically predicted the availability of a minimum energy required to switch the system from the "ON" mode into the "OFF" one. Landauer himself proceeded from the consideration of the uncertainty relation [5]. But

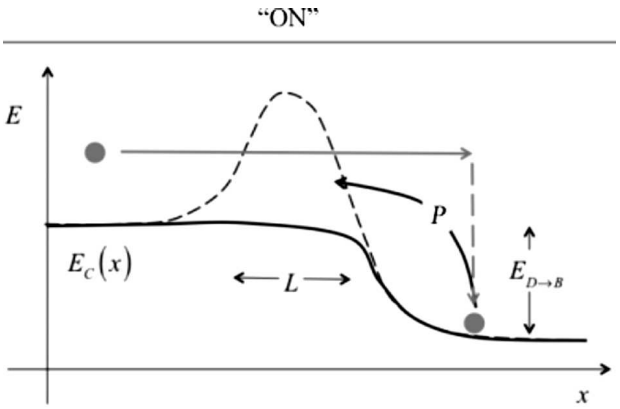

Fig. 1. Switching of the MOSFET transistor from one mode to another one: the solid curve corresponds to the "ON" mode, and the dashed one to the "OFF" mode. The switching probability equals $1 - P$, where P is the probability of thermoionic reemission from the drain to the source. Adapted from work [1]

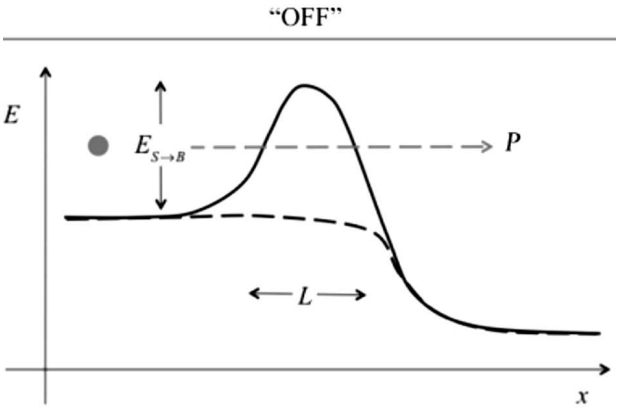

Fig. 2. In the "OFF" mode, there is a probability P of the quantum-mechanical tunneling of an electron through a barrier with the width L. Adapted from work [1]

the same result can be obtained using simple illustrative considerations of the MOSFET band structure. Figure 1 shows the energy of the conduction band bottom as a function of the x-coordinate directed along the channel of an n-channel MOSFET operating in the "ON" mode. A high positive voltage applied to the gate virtually vanishes the barrier, and electrons can pass from the source through the conduction channel to the drain. We assume that the transport in the channel is ballistic, so the electrons give their energy to the drain and relax owing to intense inelastic interactions in the drain terminal.

Even this simple model allows the minimum energy required to switch from one mode to another one to be estimated. The high voltage on the gate in the "ON" mode eliminates the barrier between the source and the channel, but the barrier $E_{D\to B}$ between the drain and the barrier top survives, because a positive voltage is applied to the drain. As the electrons have thermally relaxed in the drain and transferred their kinetic energy to lattice phonons, there remains some probability P that the electrons can still overcome the barrier $E_{D\to B}$ and return back to the source: in this case, no switch will occur. Demanding that this probability should be less than $1/2$, we obtain

$$P = e^{-E_{D\to B}/kT} < \frac{1}{2}. \tag{1}$$

From whence, we estimate the minimum switching energy

$$E_{\min} \equiv E_S|_{\min} = kT \ln 2, \tag{2}$$

which equals 0.017 eV at room temperature. This simple consideration is heuristic. However, both the Landauer pioneering work [5] and the modern detailed analysis [6] lead to the same results for the minimum switching energy.

In the framework of the same simple scheme, it is also possible to estimate [1] the minimum length of the MOSFET conduction channel. As is clear from Fig. 2, the barrier height must be at least not lower than $E_{\min}$ in the "OFF" mode, which guarantees that the electrons can overcome the barrier with a probability of less than $1/2$. The minimum barrier width (the channel length) is determined by the quantum-mechanical tunneling through the barrier. The probability that an electron from the source will tunnel through the barrier can be estimated in the Wentzel–Kramers–Brillouin (WKB) approximation (see, e.g., work [7]), which brings about a known formula for the tunneling probability of a particle with the energy E and the mass m^* through the barrier with the potential $V(x)$ between the points x_1 and x_2,

$$P \approx \exp\left(-\frac{2}{\hbar}\int\limits_{x_1}^{x_2} \sqrt{2m^*(V(x)-E)}dx\right). \tag{3}$$

Since we estimate the minimum channel length from the side of larger MOSFET channel length values, for which the quasiclassical motion approximation is valid, so we can put $\hbar \to 0$, and the WKB approximation is unequivocally applicable in this case.

From the requirement that the tunneling probability should be lower than $1/2$ in the "OFF" mode, it follows that

$$P = e^{-2\sqrt{2m^* E_{S\to B}}\, L/\hbar} < \frac{1}{2}. \tag{4}$$

In Eq. (3), the barrier potential was assumed to be rectangular along the entire channel length L. Let us put it equal to the minimum value $E_{\min}$ described by Eq. (2). Then, for the minimum length of the conduction channel, we obtain

$$L_{\min} \approx \frac{|\ln(1/2)|}{2} \frac{\hbar}{\sqrt{2m^* E_{\min}}}. \tag{5}$$

Let us estimate the value of this minimum length for a thin inverse n-channel in Si(100), where the quantization in the direction toward the substrate depth already takes place. As can be shown [8], the lower subband with $n = 1$ is characterized in this case by the effective mass in the localization direction $m_l^* = 0.92\, m_0$ and a valley degeneracy equal to 2. But the effective mass in the direction x of free motion along the channel plane is $m_t^* = 0.19\, m_0$, and it is the latter that must be taken into account in Eq. (5). As a result, we obtain the value $L_{\min} \approx 1.2$ nm, which may give an impression that there is a possibility to reduce the channel lengths in silicon MOSFETs by at least several times as compared to those that are currently used in electronics.

However, the real problems that now arise at the transistor scaling call this statement into question. It is so because, although the length of experimental channels of 3 nm was reached more than 15 years ago [3], they have not yet been implemented in practice in electronics. One of the reasons may be the fact that the simple estimate (5) was made without considering the actual form of a potential profile in the channel, actually by assuming that there is no drain voltage.

However, this potential is no longer rectangular under the condition that a direct bias is applied to the drain, but has a profile similar to that shown in Fig. 2. In this case, as is known from the modern, carefully constructed theory of nanotransistors [1], a rather narrow region with the length $\ell \ll L$ near the barrier top, where the potential at the source varies insignificantly, and the electric field is almost absent, is critical for the classical passage of the barrier by electrons [9]. An electron that managed to pass through this narrow region is drawn out by a strong electric field to the drain, even if it undergoes collisions.

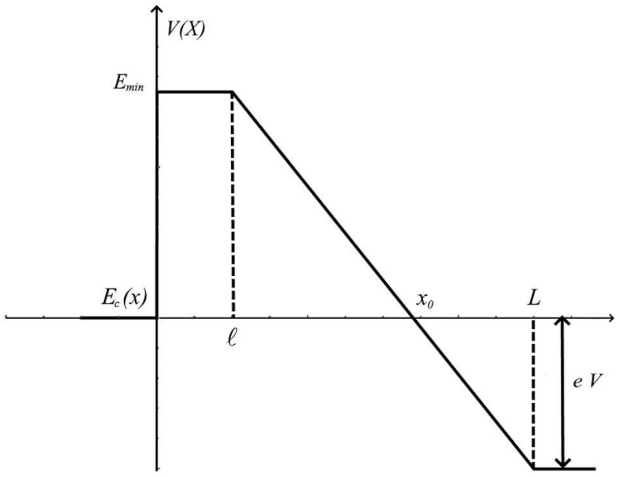

Fig. 3. Approximation of the potential profile in the MOSFET channel with the length L. Electrons tunnel through the barrier from the source (left) to the drain (right). The voltage V is applied to the drain

Therefore, the real potential corresponding to the situation where the voltage V is applied to the drain is approximated by the profile shown in Fig. 3,

$$V(x) = \begin{cases} E_{\min}, \ 0 < x \leqslant \ell, \\ E_{\min} - \dfrac{E_{\min} + |\mathrm{eV}|}{L - \ell}(x - \ell), \ \ell < x \leqslant L. \end{cases} \tag{6}$$

A similar approximation is widely used to consider the passing through MOSFETs with scattering [10]: although the specific potential profile is a result of the complicated numerical simulation, the results of such simulations are in agreement with the values obtained using Eq. (6), when "fitting" the curves at the point of virtual drain, along almost the whole length of the channel with an accuracy of 15%.

It should be noted that the length ℓ is a function of the gate voltage V, and $\ell \to L$ at $V \to 0$ (it is clear that, without a voltage, the whole channel becomes a region with a constant potential and the zero electric field). However, as was shown by numerical simulations [11], the growth of V very quickly reduces ℓ to a certain constant value. Therefore, in this wide interval of drain voltages, we can assume that $\ell = \xi L$, where the parameter $\xi \approx 0.1$.

Now, it is necessary to integrate Eq. (3) with regard for dependence (6) within the interval from 0

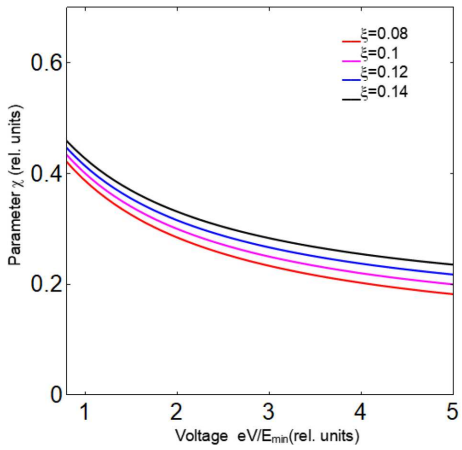

Fig. 4. Dependences of the parameter χ on $eV/E_{\min}$ for various ξ-values

to x_0 at which the lower expression in Eq. (6) equals zero, since the tunneling through the barrier shown in Fig. 3 is possible only for positive electron energy values. This integral is taken exactly and leads to the modification of result (5), namely,

$$L_{\min} \approx \frac{|\ln(1/2)|}{2} \frac{\hbar}{\sqrt{2m^* E_{\min}}} \frac{1}{\chi}, \tag{7}$$

where

$$\chi = \xi + \frac{2}{3} \frac{(1-\xi)}{1+\mathrm{eV}/E_{\min}}. \tag{8}$$

Figure 4 demonstrates the dependences of parameter (8) on the dimensionless voltage $eV/E_{\min}$ for several values of ξ, which, by the order of magnitude, correspond to a value obtained as a result of the numerical simulation [11]. As one can see, in the voltage interval, where expression (8) is valid (recall that at very low voltages, the parameter ξ is also a function of voltage), we obtain the value $\chi \sim 0.3$, which relatively weakly depends on the further voltage growth. Taken expression (7) into account, the value of the minimum channel length becomes three times as large, $L_{\min} \approx 1.2$ nm.

The results obtained above correlate with the results of simulations of the quantum transport in Si NW MOSFET [12]. As was shown in the cited work, if the length of the gate part of the channel $L_G = 12$ nm, the current I_{OFF} flows almost completely above the barrier. Such a transistor operates in the ordinary classical mode controlled by the barrier. When the channel length decreases to 10 nm, a small fraction of electrons already tunnel through the barrier, but the ordinary mode controlled by the barrier still remains mainly preserved. At $L_G = 7$ nm, a significant fraction of the electrons that form the current I_{OFF} already tunnel through the barrier. Finally, at $L_G = 5$ nm, only some fraction of the current I_{OFF} is created by the tunneling through the barrier. Such a length of the conduction channel makes it impossible to control the current by governing the barrier height, because the barrier became permeable for electrons.

The results obtained in this work are in agreement with the generally recognized idea that the classical model of the passage through the MOSFET is quite applicable to Si-based transistors with a conduction channel length of down to 10 nm and even slightly shorter. However, the further scaling to 5 nm already poses serious problems of both the applied (the growth of the role of parasitic resistances and capacitors for very short conduction channels) and fundamental characters owing to the tunneling through the barrier. The numerical simulation of transistors with stressed substrates and a special choice of their orientation with respect to the conduction channel direction (which allows the value of the effective mass in Eq. (7) to be made larger) show that, most likely, it is possible to realize the acceptable functioning mode of a MOSFET with the conduction channel length even slightly shorter than 5 nm [13]. However, it is obvious that this value cannot be significantly reduced in the future because of the fundamental limitations associated with the quantum-mechanical nature of the electron motion through such a short channel.

3. Conclusions

In this work, the minimum channel length of a MOSFET transistor, which is the main device of modern electronics, has been estimated. Considering the real potential profile in the channel, it is demonstrated that, in the presence of a drain voltage, electrons tunnel through a region that is significantly shorter than the physical length of the channel L. Therefore, the estimate of the minimum channel length in the silicon-based MOSFET, $L_{\min} \approx 1.2$ nm, which is available in the literature, becomes substantially underestimated owing to quantum restrictions. From whence, it is clear why, after reaching the working channel lengths of 5 nm, the developers did not manage to achieve the long-declared values of 3 nm pro-

vided the proper level of the transistor functionality. The estimates made in our work confirm that the fundamental limits of the scaling have almost been reached for silicon-based MOSFETs.

1. M. Lundstrom. *Fundamentals of Nanotransistors* (World Scientific, 2018).
2. B. Doris, O. Dokumaci, M. Ieong, A. Mocuta, Y. Zhang, T.S. Kanarsky. Extreme scaling with ultra-thin Si channel MOSFETs. *Digest Int. Electron Dev. Meet.* 267 (2002).
3. H. Wakabayashi, S. Yamagami, N. Ikezawa, A. Ogura, M. Narihiro. Sub-10-nm Planar-Bulk-CMOS devices using lateral junction control. *IEEE Int. J. Electron Dev. Meet.* 989 (2003).
4. Yu.O. Kruglyak, M.V. Strikha. Physics of nanotransistors: 2D electrostatics of MOS and virtual drain model. *Sensorna elektronika i mikrosystemni tekhnolohii* **16**, 19 (2019) (in Ukrainian).
5. R. Landauer. Irreversibility and heat generation in the computing process. *IBM J. Res. Develop.* **5**, 183 (1961).
6. J.D. Meindl, J.A. Davis. The fundamental limit on binary switching energy for terascale integration (TSI). *IEEE J. Solid State Circuits* **35**, 1515 (2000).
7. A.S. Davydov. *Quantum Mechanics* (Pergamon Press, 1976).
8. Yu.O. Kruglyak, M.V. Strikha. Physics of nanotransistors: Gate voltage, surface potential, and moving electron charge in a massive MOS structure and in a thin SOI. *Sensorna elektronika i mikrosystemni tekhnolohii* **16**, 5 (2019) (in Ukrainian).
9. Yu.O. Kruglyak, M.V. Strikha. Nanotransistor physics: Landauer–Datta–Lundstrom transport model and ballistic MOSFET. *Sensorna elektronika i mikrosystemni tekhnolohii* **16**, 5 (2019) (in Ukrainian).
10. Yu.O. Kruglyak, M.V. Strikha. Physics of nanotransistors: Electron scattering and MOSFET passage model. *Sensorna elektronika i mikrosystemni tekhnolohii* **17**, 16 (2020) (in Ukrainian).
11. A. Majumdar, D.A. Antoniadis. Analysis of carrier transport in short-channel MOSFETs. *IEEE Trans. Electron Dev.* **61**, 351 (2014).
12. A.D. Franklin, M. Luisier, Shu-Jen Han, G. Tulevski, C.M. Breslin, L. Gignac, M.S. Lundstrom, W. Haensch. Sub-10 nm carbon nanotube transistor. *Nano Lett.* **12**, 758 (2012).
13. R. Mehrotra, Sung Geun Kim, T. Kubis, M. Povolotskyi, M.S. Lundstrom, G. Klimeck. Engineering Nanowire n-MOSFETs at Lg $<$ 8 nm. *IEEE Trans. Electron Dev.* **60**, 2171 (2013).

Received 22.12.20.
Translated from Ukrainian by O.I. Voitenko

М.В. Стріха, А.І. Курчак

ФУНДАМЕНТАЛЬНІ ОБМЕЖЕННЯ ДЛЯ ДОВЖИНИ КАНАЛУ ПРОВІДНОСТІ MOSFET З УРАХУВАННЯМ РЕАЛЬНОГО ВИГЛЯДУ БАР'ЄРНОГО ПОТЕНЦІАЛУ

В статті оцінено мінімальну довжину каналу транзистора MOSFET, який є основним пристроєм сучасної електроніки. Врахування реального вигляду потенціалу в каналі показує, що за наявності напруги на стоку електрон тунелює крізь область, суттєво коротшу від фізичної довжини каналу L, і тому наявна в літературі оцінка мінімальної зумовленої квантовими обмеженнями довжини каналу в кремнієвому MOSFET $L_{\min} \approx 1{,}2$ нм є суттєво заниженою. Звідси зрозуміло, чому після досягнення робочих довжин каналу в 5 нм так і не вдалося вийти на вже давно декларовані значення в 3 нм при збереженні належного рівня функціональності роботи транзистора. Зроблені в нашій роботі оцінки підтверджують: фундаментальних меж масштабування кремнієвих MOSFET вже майже досягнуто.

Ключові слова: транзистор метал–діелектрик–напівпровідник, мінімальна довжина каналу, тунелювання крізь бар'єр.