Determination of the hole effective mass in thin silicon dioxide film by means of an analysis of characteristics of a MOS tunnel emitter transistor

2005 J. Phys.: Condens. Matter 17 8057

(http://iopscience.iop.org/0953-8984/17/50/023)

J. Phys.: Condens. Matter **17** (2005) 8057–8068 doi:10.1088/0953-8984/17/50/023

Determination of the hole effective mass in thin silicon dioxide film by means of an analysis of characteristics of a MOS tunnel emitter transistor

M I Vexler, S E Tyaginov and A F Shulekin

A F Ioffe Institute, 26 Polytechnicheskaya, 194021 St Petersburg, Russia

E-mail: mega@pulse.ioffe.rssi.ru

Received 16 May 2005, in final form 27 October 2005
Published 2 December 2005
Online at stacks.iop.org/JPhysCM/17/8057

Abstract
The value of $m_h = 0.33\, m_0$ has been experimentally obtained for hole effective mass in a tunnel-thin (2–3 nm) SiO_2 film. The use of this value ensures the adequate modelling of a direct-tunnelling hole current in MOS devices. For the first time, in order to determine m_h, the characteristics of a MOS tunnel emitter transistor have been mathematically processed, that allows for the precise estimation of the effective oxide thickness, as the electron effective mass in SiO_2 is independently known from the literature. The formulae for simulation of currents in a tunnel MOS structure are listed along with the necessary parameter values.

1. Introduction

Silicon dioxide (SiO_2) is one of the most important materials of contemporary microelectronics [1]. Its main application is the use as a gate insulator layer in the field-effect transistor; in the modern variants of this device, the SiO_2 film thickness may be less than 3–4 nm [2, 3]. In such a case, the gate-to-substrate cross-section constitutes, in fact, a tunnel metal–oxide–semiconductor (MOS) structure [4, chapter 9], figure 1. In a field-effect transistor, the tunnelling is a parasitic effect, but there are another elements—MOS photodiodes (e.g. [5]) and transistors with a tunnel MOS emitter [7–12]—whose operation is just based on the tunnelling through the oxide. Anyway, the charge transport through thin SiO_2 film requires careful consideration. It needs to be regarded in modelling, for which the information about the parameters of an upper and lower tunnel barrier in a MOS system is essential.

The band structure parameters of bulk SiO_2 (bandgap $E_{g,bulk} = 8.9$ eV [13], effective masses in allowed bands $m_{e,bulk} = 0.5\, m_0$, $m_{h,bulk} = 3 - 10\, m_0$ [1], band discontinuities at the Si/SiO_2 interface in thick MOS structure [4, chapter 7]), are reliably known. However, it is not at all evident whether these values are applicable to a tunnel-transparent film, containing just a few molecular layers, not to mention possible technology-related deviations. Furthermore,

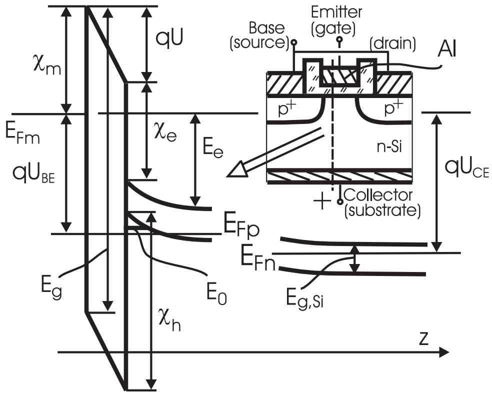

Figure 1. Band diagram of an Al/SiO_2/n-Si tunnel MOS emitter transistor and the cross-section of this transistor.

when the tunnel transport is considered, the most interesting is the effective mass (or possible effective masses) of carriers in the forbidden band of the insulator, rather far from the band edges, so that the automatic use of a mass from the nearest allowed band is questionable.

The aim of this work is to determine the effective hole mass in a thin (2–3 nm) silicon dioxide layer. For the study of tunnelling parameters, we will treat the characteristics of a bipolar tunnel MOS emitter transistor (Al/SiO_2/n-Si) as it enables the separate consideration of the co-existing electron and hole components of tunnel current. The mentioned device is topologically identical to a regular field-effect transistor and may be treated as its special connection: the gate is taken as emitter, the substrate as collector, and the source and drain are shorted together, acting as a contact to the inversion base layer [6–9]. By changing the collector and base voltages, the conditions for current transport in SiO_2 may be varied.

The attempts to find the effective masses by measuring the characteristics of MOS tunnel structures have also been undertaken before ([14–16] etc). There are three reasons for returning to this question. First, in many previous works (e.g. [14, 15]), when speaking about 'tunnelling' in a MOS structure, only the transport through the upper barrier has been considered; at the same time, the literature data on hole effective mass in a thin SiO_2 layer are very contradictory (0.28 m_0 [17], 0.34–0.37 m_0 [18], 0.51 m_0 [19]). Secondly, the progress in technology and theory of thin MOS structures can now provide better reliability of results. Furthermore, the tunnel MOS emitter transistor is used for such measurements for the first time, that warrants a certain methodical independence.

2. On the problem formulation

From a physical standpoint, the question should be formulated not on the carrier masses within a barrier, but on the dependence of the wavevector of an electron on its energy in the forbidden gap of a dielectric. Nevertheless, for the sake of simplicity, one usually speaks about the masses. These, in turn, act as fitting coefficients for the well accepted 'tunnel' formulae, ensuring the coincidence between the experiment and simulations performed within a specific model of a whole device. Also in this work, a rather simple model of a tunnel transistor (details are reported further) will be used, and two masses m_e and m_h for upper and lower SiO_2 barriers, respectively, will be considered admitting that they are different from $m_{e,bulk}$ and $m_{h,bulk}$. We therefore postulate the existence of two parabolic energy bands for the complex wavevectors.

Of course, such an approach fails to provide a smooth transition to 'thick' MOS structures and moreover to the case when the tunnel barrier transforms from the trapezoidal to the triangle one, due to the application of a high voltage. It should be noted, however, that the provision of this transition lies outside any practical needs. Indeed, if the SiO_2 film is as thin as 2–3 nm, the electron and hole transport can only occur by direct tunnelling, without entering the allowed band of the oxide, because the latter—especially for holes—would require electric fields above the breakdown for SiO_2 (10^7 V cm^{-1} [1]). As for thicker MOS structures, their use will become more and more restricted, within the general scaling tendencies in microelectronics [2].

When writing the direct-tunnelling probability through the oxide, T, the particle will be considered to simultaneously interact with both upper and lower barriers, independently of its energy:

$$T = \Theta_e + \Theta_h - \Theta_e\Theta_h \tag{1}$$

where

$$\Theta_e(E_{ze}) = T_{Re}\exp\left(-\frac{4\sqrt{2m_e}d}{3\hbar qU}\left[(qU + \chi_e - E_{ze})^{3/2} - (\chi_e - E_{ze})^{3/2}\right]\right) \tag{2}$$

$$\Theta_h(E_{zh}) = T_{Rh}\exp\left(-\frac{4\sqrt{2m_h}d}{3\hbar qU}\left[(\chi_h + E_{g,Si} + E_{zh})^{3/2} - (\chi_h + E_{g,Si} + E_{zh} - qU)^{3/2}\right]\right) \tag{3}$$

$$T_{Re(h)} = \frac{4|v_{z,Si}|\langle|v_{ze(h)}|\rangle}{|v_{z,Si}|^2 + \langle|v_{ze(h)}|\rangle^2}. \tag{4}$$

Here U (>0) is for the oxide voltage and d is SiO_2 thickness. The sense of symbols E_{ze} and E_{zh} will be explained further. In the pre-exponential factor $T_{Re(h)}$, v_z means the carrier velocity in the tunnelling direction z in Si and also in the 'complex' e- or h-zone of SiO_2, averaged over the film. Assuming a parabolic dispersion law, these velocities are easily found based on the energy components of a particle. $T_{Re(h)}$ provides $T \to 0$ near the Si band edges, and in other cases $T_{Re(h)} \approx 1$. For a poly-Si electrode, this factor would have been written in a similar form [20]. The equations (2) and (3) obtained by the Wentzel–Kramers–Brillouin (WKB) method [21, chapter 7] are well known and used (usually without alliance (1)) in a large number of works on MOS tunnel structures, although—from the fundamental viewpoint—they may be criticized [22].

3. Established tunnel barrier parameters in an Al/SiO_2/Si system

Today, it may be considered as established that the forbidden gap E_g of a tunnel-thin SiO_2 is the same as $E_{g,bulk}$. One of the latest confirmations to this, obtained by optical methods, may be found in [23]. The barrier heights χ_m (3.17 eV), χ_e (3.15 eV) and χ_h (4.63 eV, because $E_g = 8.9$ eV and $E_{g,Si} = 1.12$ eV) also became (see figure 1) almost commonly accepted. All the data presented in the modern literature are very close to these values. Anyway, the attempts to attribute the greatly reduced heights to the barriers at the heterointerface with ultrathin oxide, which were made earlier (e.g. [8]), should not be considered further.

For an upper-barrier effective carrier mass m_e, the value of 0.42 m_0 is most often encountered in the recent literature [24, 25], at least in a thickness range of 1.5–4 nm. Although it is still not unequivocally and finally established, we will take $m_e = 0.42\ m_0$ in our work. Note that the data of many publications (e.g. [26]) are close to this value; minor discrepancies may already be due to the details of the models applied by different authors. As with χ_e, the use of substantially reduced m_e [7, 15] looks unjustified. That the upper-barrier tunnelling has been studied and parametrized more carefully than the lower-barrier transport is partly because the tunnel leakage from a channel and therefore the static power consumption is critical for an n-MOSFET, not for a p-MOSFET, due to the smaller conduction band offset at the Si/SiO_2 interface.

4. Effective tunnel SiO_2 thickness in a MOS structure

The correct description of the SiO_2 film thickness is a prerequisite for determination of the tunnel barrier parameters. Except the averaged (nominal) thickness d_n, the oxide layer is characterized with the standard deviation σ_d reflecting the statistical distribution of thickness over the device area. Because the dependence of the tunnel current density on the local thickness d is rather strong, the current will crowd in the thinnest device parts. As shown in [27], a satisfactory approximation for the total current (for $\sigma_d < 4$ Å) may be obtained by substitution of the 'effective' thickness $d = d_{\mathrm{eff}}$, instead of $d = d_n$, into formulae (2) and (3): $d_{\mathrm{eff}} = d_n - 0.5\sigma_d^2$ (all values in Å).

The artificial reduction of χ_e and m_e which was often made in earlier works might arise from the disregard for the thickness non-uniformity. In order to 'justify' the flux of a larger current, that could be expected for a given d_n, measured ellipsometrically or through the capacitance–voltage (C–V) characteristics, one had to introduce lower barrier height and/or carrier mass. (While the dependence of a local current density on the thickness is, roughly, exponential, the capacitance is inversely proportional to d, so that the effect of SiO_2 thickness inhomogeneity on the capacitance is not so pronounced.)

That the reliable knowledge of d_{eff} is indeed very important, becomes evident at least from the following example. The increase of a thickness by 0.2 nm should result in a current reduction by approximately an order of magnitude; roughly the same reduction will be obtained if one puts 0.5 m_0 instead of 0.42 m_0 for the mass m_e.

5. Description of the tunnel MOS emitter transistor

The theory of a tunnel MOS emitter transistor is developed elsewhere [7–9, 12], so only the main points will be recalled here.

For an analysis of the partition of the applied collector bias U_{CE}, a simple model [12] is used. This model considers the quantum confinement effect, but all the holes in the inversion layer are assumed to be located at the sole energetic level E_0. Such an approximation is justified for the limit of a strong electric field in SiO_2, occurring in the regular operation mode of a tunnel transistor.

For simulation of an electron tunnel current, the equation [28]

$$j_e = \frac{4\pi \nu_e m_{e\perp} q}{h^3} \int_0^{+\infty} (f_m(E) - f_n(E)) \int_0^E T \, dE_\perp \, dE \tag{5}$$

is used, where for calculation of T according to equations (1)–(4) the values of

$$E_{ze} = E - E_\perp m_{e\perp} m_e^{-1} \qquad E_{zh} = E + E_\perp m_{e\perp} m_h^{-1} \tag{6}$$

are substituted. Here, $E_\perp$ is the energy associated with the particle motion in silicon in the plane of the Si/SiO_2 interface, f_m and f_n are Fermi functions for a metal and for the conduction band of Si, and $m_{e\perp}$ denotes the mass in the interface plane in the conduction band of Si. For a hole current from the inversion layer, the expression is somewhat different, due to the quantization:

$$j_h = \frac{q \nu_h m_{h\perp}}{\tau(E_0)\pi\hbar^2} \int_{-\infty}^{-E_g - E_0} (f_m(E) - f_p(E)) T \, dE. \tag{7}$$

Table 1. List of values used (except the barrier parameters).

Symbol	Parameter denoted	Value
$E_{g,Si}$	Si bandgap	1.12 eV
ε_{Si}	Si permittivity	11.9
ε	SiO_2 permittivity	3.9
t	Temperature	300 K
$P(E_e)$	Quantum yield of Auger generation	As in [30]
γ	Quantum yield of impact ionization	As in [4], chapter 1
j_{diff}	Base–collector junction current	As in [8]
ν_e	Transversal degeneracy for electrons	6
ν_h	Transversal degeneracy for holes	3
	Effective masses in Si-100 (Si-111):	
m_{ez}	Of an electron in z direction	0.432 m_0 (0.258 m_0)
$m_{e\perp}$	Of an electron in transversal plane	0.341 m_0 (0.358 m_0)
m_{hz}	Of a hole in z direction	0.260 m_0 (0.392 m_0)
$m_{h\perp}$	Of a hole in transversal plane	0.297 m_0 (0.330 m_0)

In this equation, τ is a time between sequential collisions of a hole with the border of a tunnel barrier, and for calculation of T the values

$$E_{ze} = E - (-E_g - E_0 - E)m_{h\perp}m_e^{-1} \qquad E_{zh} = E + (-E_g - E_0 - E)m_{h\perp}m_h^{-1} \tag{8}$$

are used. The sense of symbols f_p and $m_{h\perp}$ is the same as that of f_n and $m_{e\perp}$, but for the valence band. ν_e and ν_h are degeneracies. The null of the total energy E is taken at E_{c0} (figure 1), and this energy is always counted upwards. In place of d in (2) and (3), the value of d_{eff} needs to be inserted. It should be noted that formulae (5)–(8) implicitly regard the conservation of the transversal component of the carrier wavevector. Similar equations are used in a large number of works. Table 1 contains the parameter values adopted in our simulations (compared to the effective masses in SiO_2, they play a relatively minor role). Effective masses for Si are calculated on the basis of data [29].

From the tunnel currents, it is easy to arrive at the terminal device currents:

$$j_E = j_e + j_h \tag{9}$$

$$j_C = j_e M - j_{diff} \tag{10}$$

$$j_B = j_h + j_{diff} - j_e(M-1). \tag{11}$$

There is a multiplication factor M in these expressions, which accounts for Auger ionization (its quantum yield is $P(E_e)$; E_e is marked in figure 1) and impact ionization in collector (yield γ): $M = (1+P)(1+\gamma)$. The presence of $P \neq 0$ due to hot electron injection means the transition into the Auger transistor regime [6, 7] in which some interesting effects are observed, in particular the device bistability [5–7] at high voltages ($U_{BE} > 2.5$–3 V). For the scope of this work, the carrier multiplication M is of no special interest. In the last equations, there is also a component j_{diff}—this is a current of the base–collector junction.

6. Experimental samples

In this work, the samples of tunnel MOS emitter transistors with the averaged oxide thickness of $d_n \sim 2.0$ nm, are examined. Their fabrication procedure included the standard processes of CMOS technology; the oxidation of n-Si(100) wafers has been performed in dry oxygen. For the oxide films grown under similar conditions, the SiO_2 thickness deviation was estimated

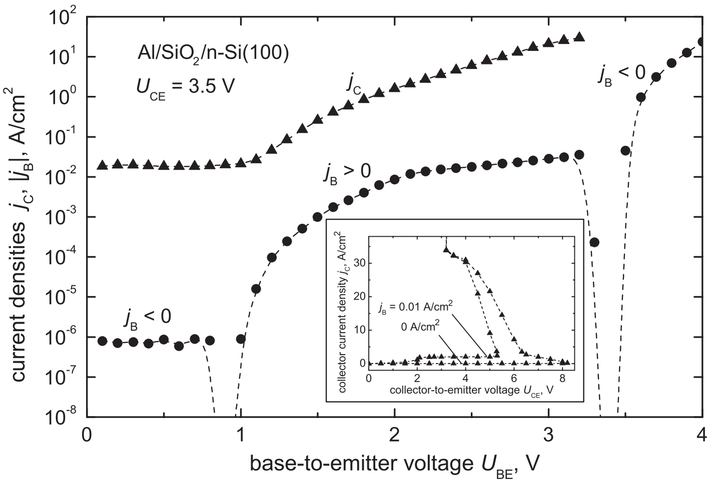

Figure 2. Experimental dependences of currents in the tunnel MOS emitter transistor on the base-to-emitter bias at fixed (rather low) collector bias; Inset—measured output characteristics. The dashed lines here and in other figures simply serve as a guide for the eye connecting the experimental points.

as $\sigma_d \sim 2$ Å using the transmission electron microscope (for a reference, the SiO_2 monolayer thickness equals 3.1 Å [1]).

At the qualitative level, the behaviour of studied devices was similar to the theoretical prediction for it and to the independent data of other authors [7–11]. The stability and reproducibility of characteristics were sufficient for reliable measurements. An example of experimental curves ($N_d = 10^{16}$ cm^{-3}) is presented in figure 2. Under the measured current densities, their averaged values (the corresponding current divided to the emitter area) are understood. Compared to our earlier work [6], the parasitic leakages in the transistor were reduced, which resulted in the gain improvement in the low-current regimes.

7. The procedure of determination of hole effective mass in SiO_2

Tunnel electron current j_e is known to greatly exceed the hole current j_h under all working conditions of the device. The collector current at $U_{CE} \geqslant U_{BE}$, as is clear from equation (10), practically equals j_e, because M is always close to unity (for the electron energy $E_e \sim 2$ eV, the quantum yield of Auger ionization is about 0.01 [30, 31]). Further, during the tunnelling between the metal and the conduction band of Si, the electrons hardly feel the lower barrier. For this reason, disposing at the measured collector current j_C and at the parameters of a tunnel barrier for electrons, including the effective mass $m_e = 0.42\, m_0$, one can estimate the effective oxide thickness d_{eff} for a given device.

The hole component j_h may be accessed through the base current j_B. If we restrict ourselves to the base voltages below 2.2 V [6] and also to low U_{CE}, the contribution of the term $j_e(M-1)$ will be unimportant. Besides, it is necessary to exclude the range of $U_{CE} < 1$ V, $U_{BE} < 1$ V and $U_{CE} < U_{BE}$ from consideration, as in that range the component j_{diff} and parasitic leakages may come into play. In this case, the base current is practically equal to the hole component of the emitter current, that enables us to find the value of m_h using the tunnel formulae and knowing d_{eff}. It is also possible to analyse the small-signal current gain $\beta_d = \mathrm{d}j_C/\mathrm{d}j_B$, instead of j_B—the smaller m_h is (at fixed m_e), the lower is the simulated gain.

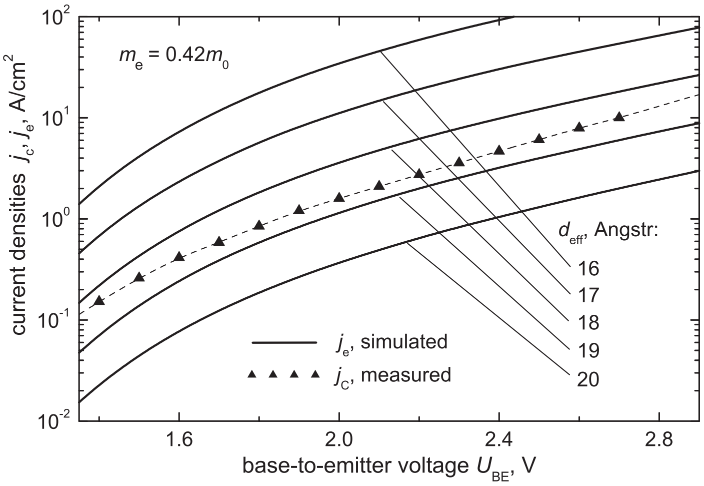

Figure 3. Illustration of the estimation of the effective oxide thickness d_{eff}—a fragment of the experimental dependence of the collector current j_C on the base-to-emitter bias U_{BE}, onto which the series of simulated dependences of an electron tunnel current j_e for different d_{eff} is superimposed. For calculation, the value of m_h is unimportant if it is larger than 0.23–0.25 m_0. In the presented range $j_C \approx j_e$.

As a matter of fact, however, there are no principal reasons for avoiding Auger ionization, but if one does not then it is better to work further from the threshold $E_e = E_{g,Si}$, where the data on quantum yield $P(E_e)$ are more reliable [31]. In particular, it may be interesting to adjust m_h so as to exactly reproduce the voltage U_{BE} corresponding to $j_B = 0$ (under this condition, usually, $U_{BE} \sim 3$–3.5 V [6, 10, 32] and $E_e \sim 2$–2.5 eV).

8. Treatment of experimental data. Result

Figures 3 and 4 show the fragments of the curves $j_B(U_{BE})$ and $j_C(U_{BE})$ selected for an analysis. These are extracted from figure 2 and accomplished with simulated lines. As one can see (figure 3), the effective SiO_2 thickness in this sample equals 18.5 Å. Such a value of d_{eff} is just used while calculating the dependences for figure 4; this allows for determining the effective hole mass m_h.

Looking at the main figure 4, we may claim that the best coincidence between the model results and measured data may be achieved if $m_h = 0.33\ m_0$ is substituted into the equations for tunnel currents (1)–(4). At the same time, it is necessary to put a slightly larger mass (~0.36 m_0), in order to correctly simulate the base voltage corresponding to $j_B = 0$; see the inset to figure 4.

No doubt, the determination of m_h based on the data for $U_{BE} = 1$–2 V is more straightforward, in the framework of our method. Indeed, even for $U_{BE} \sim 2.8$–3.2 V there remains some discrepancy in the quantum yield of Auger ionization by 1.5–3-fold between different literature sources (this may explain also the systematic shift at $U_{BE} = 2.8$–3.2 V). Furthermore, for $d_{eff} = 18.5$ Å, the range where the base current j_B changes sign (figure 2) corresponds to the electric fields in SiO_2 close to breakdown. Note here that, for the field-effect transistors, the range of $U_{BE} = 1$–2 V is much more important. However, one should say that some increase of m_h with increasing field in oxide is not impossible, and even quite probable (see also [18]).

So, the determined value of the hole effective mass in a thin SiO_2 layer is $m_h = 0.33\ m_0$.

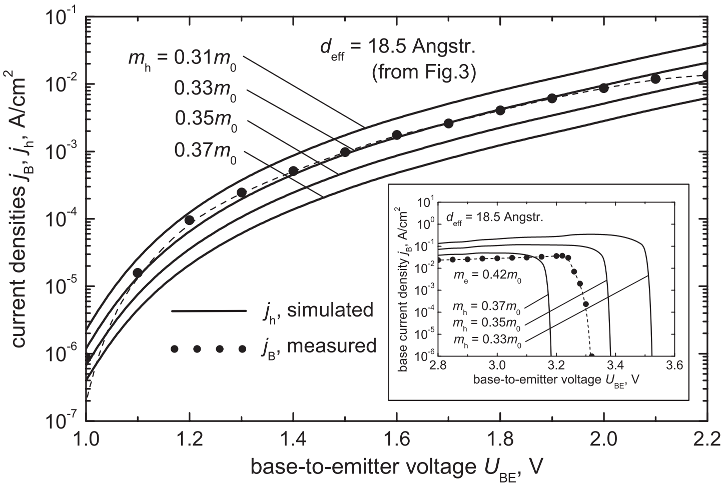

Figure 4. Illustration of the determination of the hole effective mass m_h in the oxide—a fragment of measured dependence of the base current j_B on the bias U_{BE} with the superimposed series of simulated dependences of a hole tunnel current j_h for different m_h. In the range of the main plot, $j_B \approx j_h$. Inset—simulated and measured dependences of $j_B(U_{BE})$: satisfaction of the condition $j_B = 0$ through the adjustment of the m_h value.

9. Comparison to other data and to other methods

The obtained parameter m_h, like the values of m_h from some works of other authors [16–19], is much less than the effective hole mass in 'thick' SiO_2. Our result is especially close to the data of [18] (0.34–0.37 m_0); minor differences between m_h may arise from the general differences in models of a MOS tunnel structure, as has already been mentioned in respect of m_e.

An advantage of using the MOS tunnel emitter transistor as a tool for studying m_h consists in the possibility of reliable estimation of the oxide thickness in the same device, for which the measurements of m_h are made, and by the same methods (current–voltage characteristics). The authors used the MOS structures on n-Si (p-MOSFETs) with the p^+-poly-Si electrode for the determination of m_h, and had either to achieve the C–V measurements for thickness estimation or to extract the thickness from I–V curves of similar n-MOSFETs fabricated in the same cycle, that worsens the accuracy.

Although the value $m_e = 0.42\, m_0$ adopted in our work seems to be quite reliable, it might be not useless to mention that the variations of m_e within the range of $0.3 \ldots 0.5\, m_0$ (this is the maximal spread ever faced with) would have resulted in variations of our estimated thickness d_{eff} from 22 down to 17 Å, yielding changes of m_h within $0.22 \ldots 0.40\, m_0$. Related to this, it is interesting that usually, if the masses $m_e \neq 0.42\, m_0$ and m_h are reported together somewhere, they are both larger or both smaller than our values, e.g. a pair 0.5 m_0, 0.43 m_0 [33], so that the 'current gain' remains roughly like that in our treatment. As for the too large m_e, m_h, one may admit that at least in some corresponding works the near-surface quantization in silicon was ignored. Hence, the oxide voltage (for the given terminal bias) was overestimated and it was necessary to introduce larger masses.

In general, the availability of data on current gain of the transistors enables a more thorough analysis of some details. In this connection, we now consider one common question related to the model of tunnel transport of carriers in silicon dioxide film.

In some works [7, 8, 34], the lower barrier in SiO_2 has been treated as impermeable, so that the tunnelling particles always interact only with the upper barrier. This is the so-called

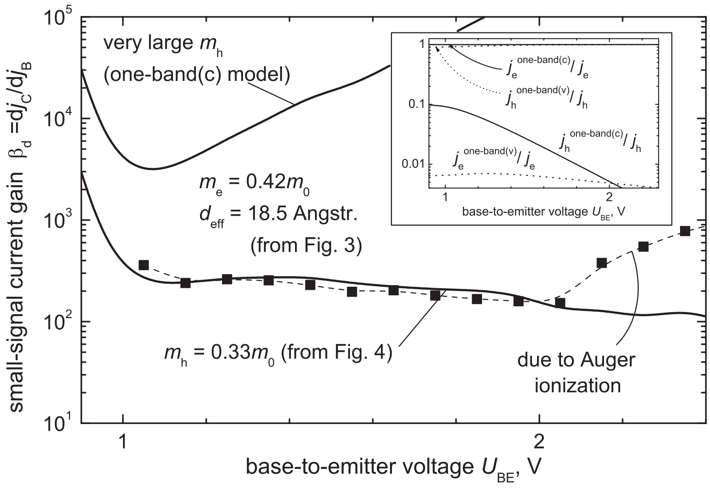

Figure 5. Experimental dependence of the small-signal current gain β_d on the base-to-emitter bias U_{BE} and its model approximation using the obtained value of a hole effective mass m_h. The discrepancy for $U_{BE} > 2.2$ V is due to the disregard for Auger ionization in calculation of a curve for this figure. The additional curve marked $m_h = \infty$ elucidates the inapplicability of the one-band (c) SiO_2 model. Inset—the relative errors in currents j_e and j_h within the one-band (c) and one-band (v) oxide models.

one-band (c) oxide model (c = conduction). In our simulations, it is possible to replicate such a situation by artificially attributing some very large value to m_h (figure 5). However, in this approach, the current j_h is greatly underestimated (inset to figure 5) and we come to too high values of current gain (e.g. $\beta_d > 10^5$ for $d_{eff} = 18.5$ Å and $U_{BE} \sim 2$ V), which were never observed. The best values of β_d ever measured for a tunnel MOS emitter transistor—in regimes with $M = 1$—are about 500 [10]. Furthermore, if only the upper barrier had worked, the gain β_d would have increased with U_{BE} (while the opposite tendency is usually observed; only the activation of Auger ionization [6, 7] suppresses the decrease of gain; see e.g. figure 5).

One can also consider the one-band (v) oxide model presuming the interaction of carriers only with the lower barrier (v = valence). For the transport between the conduction band of Si and the metal (component j_e), this model is, of course, not applicable, but for the tunnelling from the valence band of silicon (j_h) the error is seen to lie within just ten percent (inset to figure 5). This hints that, except for low U_{BE}, we could ignore the alliance of two barriers (equation (1)) and associate the upper/lower barriers, correspondingly, with the tunnelling of electrons/holes. Such an approach is used very often. However, we still believe that it is better to regard the simultaneous interaction of a particle with two barriers, which may be done quite easily and improves the accuracy, to some extent. An alliance of two barriers will emphasize the fact that the electron tunnelling in *any* case may be interpreted as the tunnelling of a hole with the same energy in the opposite direction (and vice versa).

Commenting on the novelty of the proposed method for determining m_h, it is worth mentioning the so-called carrier separation experiment in a field-effect transistor [16, 32, 35], which lies, ideologically, close to our approach. In this experiment, the source, drain and substrate are connected together[1], and the currents of source/drain and substrate are measured.

[1] Note that the curves in figures 3, 4 are practically identical to the curves $j_B(U_{BE})$ and $j_C(U_{BE})$ corresponding to the condition of $U_{BC} = 0$. For moderate N_d, the change of U_{BC} from zero to several volts does not affect the insulator voltage U (which is created by the inversion layer charge) and the impact ionization is insignificant.

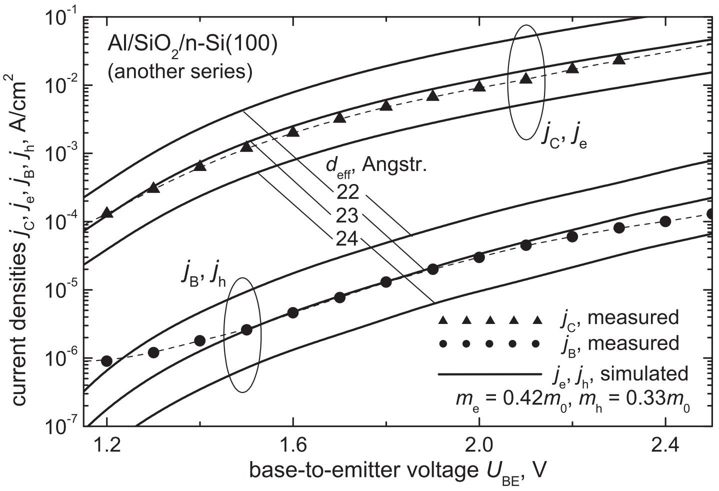

Figure 6. Estimation of the insulator thickness in a MOS tunnel emitter transistor from a series other than that used for figures 2–5. The coincidence of d_{eff} values obtained on the basis of the j_C and j_B curves speaks for an adequacy of the determined parameter m_h and of the model in general.

In fact, the field-effect transistor in this case is connected in 'our' variant, i.e. as a bipolar device, although this was not said directly (except in [35]). Such experiments were performed, however, without any relation to the determination of an effective mass; for example, the purpose of work [32] was to study the quantum yield of Auger ionization ($P = -j_B/j_C$ at high U_{BE}, in our notation). Regrettably, the data of [32, 35] are not applicable for exercising our method of determination of m_h, because they are obtained for too thick MOS structures. The carrier separation data from [16] also cannot be processed, because the p^+-poly-Si electrodes were used in that work; in such a case, the suggestion of negligible role of a lower barrier is not valid for tunnelling from the poly-Si valence band into the Si conduction band (while such a suggestion is essential for estimation of d_{eff}).

10. Additional testing of a MOS structure model

Striving to additionally verify our model with the determined parameter m_h, we tried to apply it to the extraction of oxide thickness in our sample of a MOS tunnel emitter transistor from another series (figure 6). The agreement between the values of d_{eff} obtained from the measured dependences $j_C(U_{BE})$ and $j_B(U_{BE})$ (in both cases $d_{\text{eff}} \sim 23$ Å) evidences the adequacy of the model and of the mass $m_h = 0.33\,m_0$. Note that the curve $j_B(U_{BE})$ may be used for estimations only in the range of $U_{BE} \sim 1.5 \ldots 2.2$ V, since $j_B \neq j_h$ at $U_{BE} > 2.2$ V (due to Auger ionization, as in figures 4 and 5) and also at low U_{BE} (due to the excessive parasitic current).

Further, a successful reproduction of experimental data independently obtained by other researchers could serve as one more important test. For an example, we selected the characteristics of the diode structures p^+-poly-Si/SiO_2/p-Si published in [17] (figure 7). Within the range of low voltages, in which the data are measured, the main current flows between the valence band of Si and valence band of poly-Si, so that the correct value of m_h is very critical. In modelling, the quantization of carrier motion was regarded both in substrate and in p^+-poly-Si, which, depending on the polarity of bias V, fall into the inversion or accumulation regime; for accumulation, the electrostatic model [36] was used. The expressions for a current

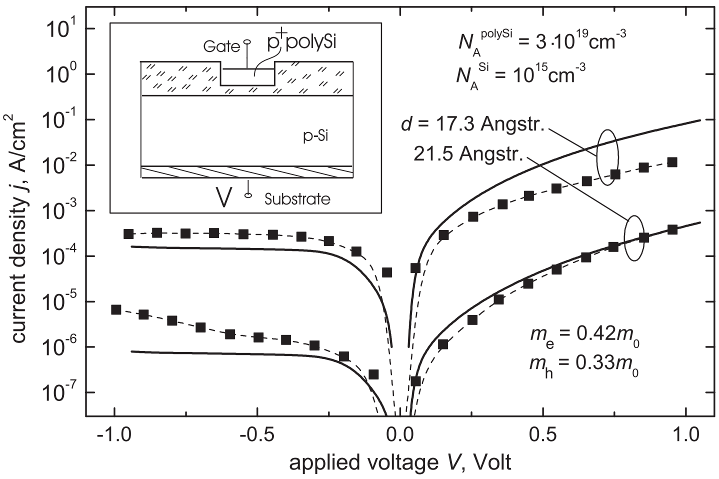

Figure 7. An independent test for correctness: the reproduction of experimental data [17] using the model from this work with the determined parameter $m_h = 0.33\ m_0$. Inset—a two-electrode p^+-poly-Si/SiO_2 /p-Si structure for which the data [17] are obtained.

from the inversion/accumulation layer are written similarly to equation (7). Figure 7 shows that the model with $m_h = 0.33\,m_0$ allows for achieving a satisfactory agreement to the experimental results[2], although the structure used for testing of our model is in all respects different from the tunnel transistor used for determining the value of m_h.

11. Conclusion

In this work, the value of a hole effective mass in thin (2–3 nm) silicon dioxide film has been experimentally determined: $m_h = 0.33\ m_0$. This value plays the role of the fitting coefficient to be substituted into the well accepted formulae for the direct-tunnelling current. It may be used in modelling the MOS devices, in particular the modern p-channel field-effect transistors, together with the previously established electron effective mass $m_e = 0.42\ m_0$. The obtained results solidify information about the properties of the technically important MOS tunnel structure. For the first time, the MOS tunnel emitter transistor has been used as a testing tool; the elaborated method may further also be used for MOS structures with alternative dielectrics.

The set of formulae recommended for the modelling is also presented in this work, as some adjustments of m_h value are possible (probably within several hundreds) depending on the details of the model adopted. In any case, one can claim that the parameter m_h responsible for the description of tunnel charge transport through lower barrier is much less than the bulk SiO_2 hole mass, which is in accordance with the recent results of other authors. Tunnelling probability was considered—for any carrier energy—as the combined probability of tunnelling via upper and lower barriers in SiO_2.

[2] Small differences in the form of curves in figure 7 may be due to the deficiencies of studied samples. So, the experimental curves for two thicknesses are not quite parallel for $V > 0$ and the device with $d = 21.5$ Å does not demonstrate the horizontal plateau at $V < 0$ as should be expected for this type of structures ([37, 4], chapter 9, [22] etc). Of course, it would be possible to find literature data which our model reproduces slightly worse or slightly better, but a thorough comparison of this kind, involving many different data, lies beyond the scope of the present work.

Acknowledgments

The authors are grateful to the Programme for Supporting the Scientific Schools of Russia, to the Russian Foundation for Basic Research, and to the Russian Science Support Foundation for financial contribution toward this work.

References

[1] Baraban A P, Bulavinov V V and Konorov P P 1988 *Electronics of SiO_2 Layers on Silicon* (Leningrad: Izd. LGU) (in Russian)
[2] 2004 International Technology Roadmap for Semiconductors http://public.itrs.net
[3] Momose H S, Ono M, Yoshitomi T, Ohguro T, Nakamura S, Saito M and Iwai H 1996 *IEEE Trans. Electron Devices* **43** 1233–42
[4] Sze S M 1981 *Physics of Semiconductor Devices* (New York: Wiley)
[5] Lai S K, Dressendorfer P V, Ma T P and Barker R C 1981 *Appl. Phys. Lett.* **38** 41–4
[6] Grekhov I V, Shulekin A F and Vexler M I 1995 *Solid-State Electron.* **38** 1533–41
[7] Ostroumova E V and Rogachev A A 1999 *Semiconductors* **33** 1027–9 (translated from Russian)
[8] Chu K M and Pulfrey D L 1988 *IEEE Trans. Electron Devices* **35** 188–94
[9] Simmons J G and Taylor G W 1986 *Solid-State Electron.* **29** 287–303
[10] Aderstedt E, Medugorac I and Lundgren P 2002 *Solid-State Electron.* **46** 497–500
[11] Yoshimoto T and Suzuki K 1993 *Japan. J. Appl. Phys.* **32** L180–2
[12] Shulekin A F, Vexler M I and Zimmermann H 1999 *Semicond. Sci. Technol.* **14** 470–7
[13] Laughlin R B 1980 *Phys. Rev.* B **22** 3021–9
[14] Depas M, Vermeire B, Mertens P W, van Meirhaeghe R L and Heyns M M 1995 *Solid-State Electron.* **38** 1465–71
[15] Brar B, Wilk G D and Seabaugh A C 1996 *Appl. Phys. Lett.* **69** 2728–30
[16] Hou Y T, Li M F, Jin Y and Lai W H 2002 *J. Appl. Phys.* **91** 258–64
[17] Ancona M G, Yu Z, Dutton R W, Vande Voorde P J, Cao M and Vook D 2000 *IEEE Trans. Electron Devices* **47** 2310–9
[18] Haque A and Alam K 2002 *Appl. Phys. Lett.* **81** 667–9
[19] Yang K-N, Huang H-T, Chang M-C, Chu C-M, Chen Y-S, Chen M-J, Lin Y-M, Yu M-C, Jang S M, Yu C H and Liang M S 2000 *IEEE Trans. Electron Devices* **47** 2161–6
[20] Register L F, Rosenbaum E and Yang K 1999 *Appl. Phys. Lett.* **74** 457–9
[21] Landau L D and Lifshitz E M 1977 *Quantum Mechanics: Non-Relativistic Theory* (*Course of Theoretical Physics* vol 3) (Oxford: Pergamon)
[22] Faigon A and Campabadal F 1996 *Solid-State Electron.* **39** 251–60
[23] Ohta A, Yamaoka M and Miyazaki S 2004 *Microelectron. Eng.* **72** 154–9
[24] Schenk A and Heiser G 1997 *J. Appl. Phys.* **81** 7900–8
[25] Ludeke R 2000 *IBM J. Res. Dev.* **44** 517–34
[26] Yeo Y-C, King T-J and Hu C 2002 *Appl. Phys. Lett.* **81** 2091–3
[27] Khlil R, El Hdiy A, Shulekin A F, Tyaginov S E and Vexler M I 2004 *Microelectron. Reliab.* **44** 543–6
[28] Vogel E M, Ahmed K Z, Hornung B, Henson W K, McLarty P K, Lucovsky G, Hauser J R and Wortman J J 1998 *IEEE Trans. Electron Devices* **45** 1350–5
[29] Moglestue C 1986 *J. Appl. Phys.* **59** 3175–83
[30] Drummond W E and Moll J L 1971 *J. Appl. Phys.* **42** 5556–62
[31] Cartier E, Fischetti M V, Eklund E A and McFeely F R 1993 *Appl. Phys. Lett.* **62** 3339–41
[32] Chang C, Hu C and Brodersen R W 1985 *J. Appl. Phys.* **57** 302–9
[33] Gritsenko V A, Nasyrov K A, Novikov Yu N and Aseev A L 2003 *Russian Microelectron.* **32** 69–74 (translated from Russian)
[34] O'Neill A G 1986 *Solid-State Electron.* **29** 305–10
[35] van der Berg M R, Nanver L K, de Boer C R, Visser C C G and Slotboom J W 2000 *Proc. SAFE-2000* (Veldhoven, Netherlands) (Utrecht: STW Technology Foundation) pp 11–4
[36] Vexler M I 2003 *Solid-State Electron.* **47** 1283–7
[37] Ghetti A, Liu C-T, Mastrapasqua M and Sangiorgi E 2000 *Solid-State Electron.* **44** 1523–31