Осма национална конференција со меѓународно учество - ЕТАИ 2007
Eighth National Conference with International Participation - ETAI 2007
Охрид, Република Македонија - Ohrid, Republic of Macedonia, 19 - 21. IX 2007

E1-4

CONDUCTION BAND IN SILICON: NUMERICAL VERSUS ANALYTICAL TWO-BAND k·p MODEL

Viktor Sverdlov, Hans Kosina, and Siegfried Selberherr
Institute for Microelectronics, Technische Universität Wien
Gusshausstrasse 27–29, A-1040 Vienna, Austria
Email: {sverdlov|kosina|selberherr}@iue.tuwien.ac.at

***Abstract*— A two-band k·p model for the conduction band of silicon is proposed and compared with other band structure models, notably the nonlocal empirical pseudo-potential method and the $sp^3d^5s^*$ nearest-neighbor tight-binding model. The two-band k·p model is demonstrated to predict results consistent with the empirical pseudo-potential method, and to accurately describe the band structure around the valley minima, including the effective masses and the band non-parabolicity. The tight-binding model, on the other hand, overestimates the gap between the two lowest conduction bands at the valley minima, which results in an underestimation of the non-parabolicity effects. When strain effects are included, the two-band k·p model gives analytical expressions for the strain-dependence of band structure parameters. Shear strain leads to profound changes in the conduction band causing a shift of the valley minima in momentum space and a modification of the effective masses. Also in the strained case, predictions of the two-band k·p model are in good agreement with those of the pseudo-potential method.**

I. INTRODUCTION

The **k·p** theory allows to describe the band structure analytically. After the pioneering work by Luttinger and Kohn [1] the six-band **k·p** method has become widely used to model the valence band in silicon. The conduction band in silicon is usually approximated by three pairs of equivalent minima located near the X-points of the Brillouin zone. It is commonly assumed that close to the minima the electron dispersion is well described by the effective mass approximation. The non-parabolicity parameter α is introduced to describe deviations in the density of states from the purely parabolic expression, which become pronounced at higher electron energies. In ultra-thin body (UTB) FETs, however, the band non-parabolicity affects the subband energies substantially, and it was recently indicated that anisotropic, direction-dependent non-parabolicity could explain a peculiar mobility behavior at high carrier concentrations in a FET with (110) UTB orientation [2]. Therefore, a more refined description of the conduction band minima beyond the usual single-band non-parabolic approximation is needed. Another reason to challenge this standard approximation is its inability to address properly the band structure modification under stress.

In biaxially stressed devices the electron mobility can be nearly doubled [3]. The reason for the mobility enhancement lies in the stress-induced band structure modification. The degeneracy between the six equivalent valleys is lifted due to stress-induced valley shifts. This reduces inter-valley scattering. In case of tensile biaxial stress applied in the $\{100\}$ plane the four in-plane valleys move up in energy and become de-populated. The two populated out-of-plane valleys have favorable conductivity masses, which together with reduced inter-valley scattering results in the observed mobility increase [4]. Biaxial stress is introduced by growing Si epitaxially on SiGe. This method, however, would require a substantial modification of the CMOS fabrication process and is not used in mass production. Instead, the semiconductor industry is exploiting techniques compatible with existing CMOS process technology. Stress in the channel is created by local stressors and/or additional cap layers. Although already successfully used in mass production, the technologically relevant case of stress along $[110]$ has received little attention within the research community. It was systematically investigated experimentally just recently [5]. Inherent to $[110]$ uniaxial stress is a shear distortion of the Si crystal lattice, which induces pronounced modifications in the conduction band. The degeneracy between the six equivalent valleys is lifted. Contrary to $[100]$ uniaxial stress, the electron mobility data for $[110]$ stress suggest that the conductivity mass depends on stress. This conclusion is also supported by recent results of pseudo-potential band structure calculations [5], [6]. Shear strain also modifies substantially both the longitudinal [7], [8] and transversal [5], [7], [8], [9] effective masses. The transversal mass determines the mobility in a FET with ultra-thin Si body. In such FETs the electron mobility enhancement induced by $[110]$ tensile stress is therefore solely due to a decrease of the conductivity mass in the stress direction [5], [7], [8]. Any dependence of the effective masses on stress is neglected within the single-band description of the conduction band and can only be introduced phenomenologically. In order to describe the dependence of the effective mass on stress a single-band description is not sufficient, and coupling to other bands has to be taken into account.

Recently, a 30 bands **k·p** theory was introduced [10]. Although universal, it cannot provide an explicit analytical solution for the energy dispersion. In this work we present an efficient two-band **k·p** theory. By comparing our results with predictions of the empirical pseudo-potential method we demonstrate that the theory accurately describes both the non-parabolicty effects and the stress induced band structure modification for general stress conditions. It accurately reproduces the stress dependence of the effective mass and of the non-parabolicity parameter. The analytical two-band **k·p** model allows to study the influence of the conduction band structure on transport properties of stressed Si.

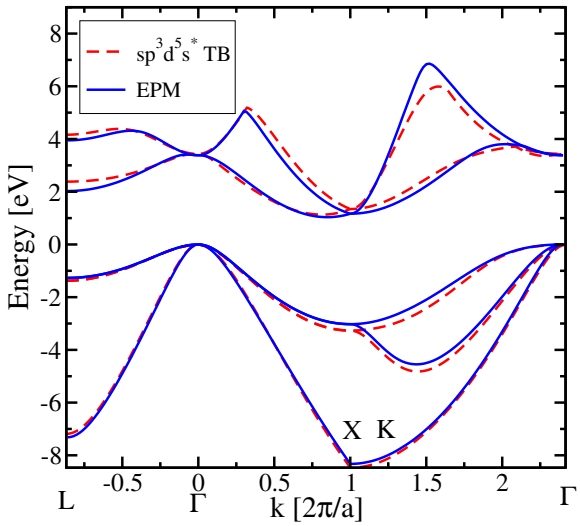

Fig. 1. Band structure from the EPM (solid) and from the $sp^3d^5s^*$ model (dashed).

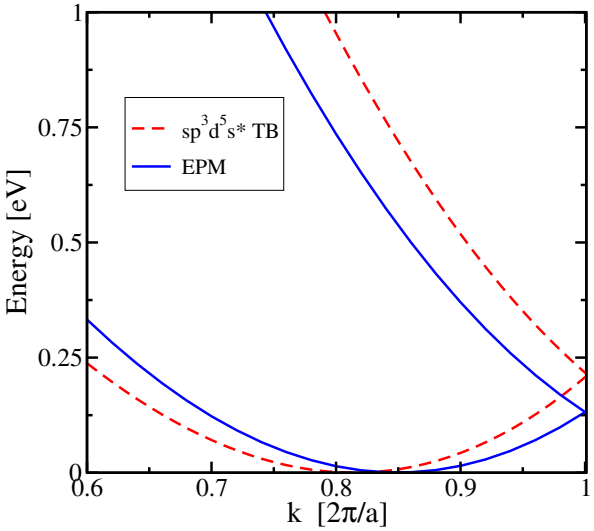

Fig. 2. Conduction bands close to the valley minimum. The EPM (solid lines) reproduces the band structure accurately.

II. Empirical nonlocal pseudo-potential and tight-binding band structure calculations

We use the empirical nonlocal pseudo-potential method (EPM) for numerical band structure calculations. The parameters of the EPM method are adjusted in order to reproduce the measurable quantities of semiconductors related to the band structure: energy gap and effective masses. The method includes spin-orbit coupling. In our calculations of the silicon band structure we used the parameters from the work [11].

Recently, empirical tight-binding methods for band structure calculations became popular. Already the sp^3s^* tight-binding model allows to reproduce reasonably well the valence band structure of silicon [12]. Recent development and calibration of a more sophisticated $sp^3d^5s^*$ model [13] has improved the reproducibility of the silicon conduction band. Agreement between the band structures obtained from the EPM and the $sp^3d^5s^*$ model with the parameters from [13] is good as shown in Fig. 1. However, the conduction band minimum in the $sp^3d^5s^*$ model is further away from the X point than in EPM, where the valley minimum is located at the distance $k_0 = 0.15\frac{2\pi}{a}$ from the X point ($0.85\frac{2\pi}{a}$ from the Γ point). This leads to an almost two times higher gap between the two lowest conduction bands at the valley minimua (Fig. 2) compared to $\Delta = 0.53$ eV found from the EPM. The increase in the gap Δ reduces the coupling between the conduction bands. Since the non-parabolicity of the lowest conduction band is determined by the coupling with other bands, as shown below, the higher gap predicted by the tight-binding method results in a substantially lower band non-parabolicty. It also results in a different shape of the constant energy lines in the k_x, k_y plane at $k_z = 0.85\frac{2\pi}{a}$. Fig. 3 shows that the EPM gives a more pronounced band warping than the $sp^3d^5s^*$ tight-binding model does.

III. Two-band **k**·**p** model

We consider the valley pair along the [001] direction. Other valleys can be analyzed in a similar fashion. The band closest to the first conduction band Δ_1 ($i = 1$) is the second conduction band $\Delta_{2'}$ ($i = 2$). The two bands become degenerate exactly at the X points. Since the minimum of the conduction band is only $k_0 = 0.15\frac{2\pi}{a}$ away from the X point, the dispersion around the minimum is well described by degenerate perturbation theory, which includes only these two bands. Diagonal elements of the Hamiltonian H_{ii} at the X point are:

$$H^0_{ii}(\mathbf{k}) = (-1)^{i-1}\frac{\hbar}{m_0}k_z p + \frac{\hbar^2 k_z^2}{2m_l} + \frac{\hbar^2 k_x^2}{2m_t} + \frac{\hbar^2 k_y^2}{2m_t}, \quad (1)$$

where $i = 1, 2$, m_0 is the free electron mass, m_t is the transversal and m_l the longitudinal effective mass. The values of k_z are counted from the X point and are thus negative. The coupling between the two bands is described by the off-diagonal terms [9]:

$$H^0_{12}(\mathbf{k}) = \frac{\hbar^2 k_x k_y}{M}. \quad (2)$$

The parameter M is obtained from the **k**·**p** perturbation theory:

$$\frac{1}{M} = \frac{1}{m_0}\left|\sum_{l\neq 1,2}(p_y)_{1l}(p_z)_{l2}\left(\frac{1}{E_l - E_1} + \frac{1}{E_l - E_2}\right)\right|.$$

We use for M the value computed by the EPM at the point $k_z = -k_0$, where the numerical value is close (but not equal) to $M \approx m_t/(1 - m_t/m_0)$ reported in [9]. With degenerate perturbation theory we obtain the following dispersion relation

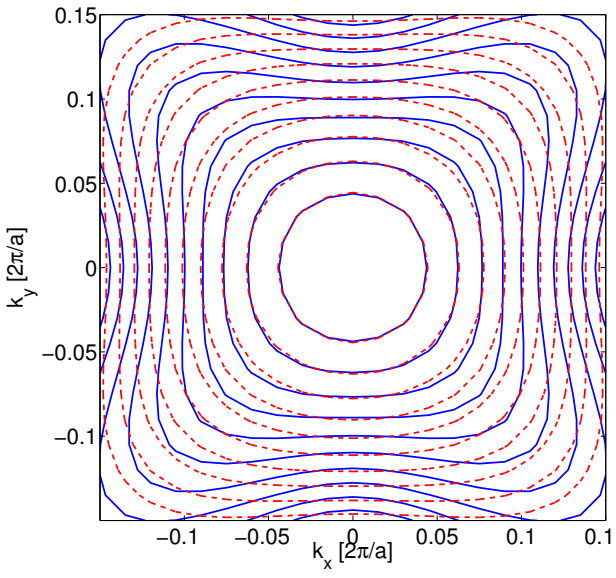

Fig. 3. Comparison between the k_x, k_y energy dispersion relations at the minimum k_0. The contour lines are spaced every 50 meV. Soild lines correspond to the EPM and the dashed lines to the $sp^3d^5s^*$ model.

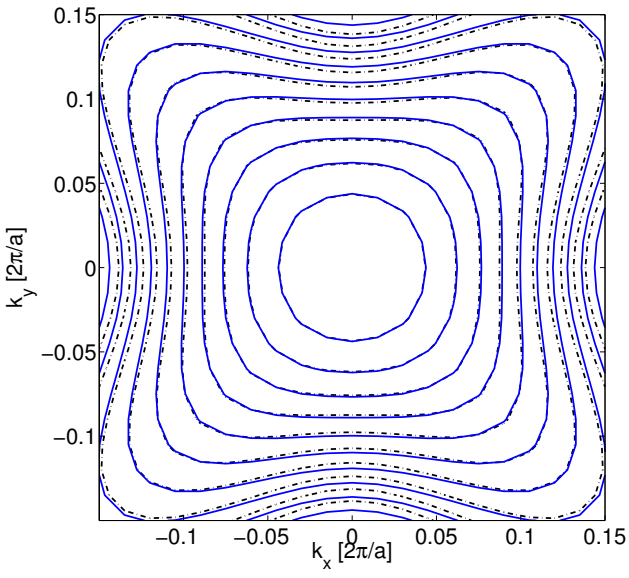

Fig. 4. Comparison of the dispersion relation (3) at the valley minimum (dashed-dotted contour lines) with the EPM results (solid lines). The distance between the equi-energy contour lines is 50 meV.

for the lowest band:

$$E(\mathbf{k}) = \frac{\hbar^2 k_z^2}{2m_l} + \frac{\hbar^2(k_x^2+k_y^2)}{2m_t} - \sqrt{\left(\frac{\hbar}{m_0}k_z p\right)^2 + \left(\frac{\hbar^2 k_x k_y}{M}\right)^2}. \quad (3)$$

Expanding (3) around the minimum $k_z = -k_0$ with respect to $q_z = k_z - k_0$, one obtains a simplified dispersion relation [14]:

$$E(\mathbf{k}) = \frac{\hbar^2 q_z^2}{2m_l} + \frac{\hbar^2(k_x^2+k_y^2)}{2m_t} - \frac{\left(\hbar^2 k_x k_y\right)^2}{M^2\Delta} - \frac{\Delta}{4}, \quad (4)$$

where $\Delta = 2\hbar k_0 p/m_0 = 2\hbar^2 k_0^2/m_l$ is the gap between the Δ_1 and the $\Delta_{2'}$ conduction bands at $k_z = -k_0$. To estimate the value of the non-parabolicity parameter, we follow [14], Appendix B, and average out the angular dependence in (4). Assuming $\alpha E(\mathbf{k})$ to be small and recasting terms, one finally obtains:

$$\frac{\hbar^2(k_x^2+k_y^2)}{2m_t} = E(\mathbf{k})\left(1+\alpha E(\mathbf{k})\right),$$

with

$$\alpha = \frac{1}{2\Delta}\left(\frac{m_t}{M}\right)^2. \quad (5)$$

Substituting the parameter values for Si in (5), we estimate $\alpha = 0.6$ eV^{-1}, which is close to the empirical value of $\alpha = 0.5$ eV^{-1}.

In Fig. 4 the analytical expression is compared to the numerical band structure obtained from the EPM at $k_z = -k_0$. Excellent agreement is found up to an energy of 0.5 eV. Fig. 4 demonstrates strong anisotropy which results in the pronounced warping of the conduction band at higher energies. Fig. 5 shows that the $sp^3d^5s^*$ model predicts less anisotropy (dashed contour lines). As indicated in Fig. 2, the gap between the two lowest conduction bands predicted by the $sp^3d^5s^*$ model is unrealistically large. This results in a smaller coupling between the bands. The solid contour lines shown in Fig. 5 obtained from (3) with an unrealistic value of Δ=1.2 eV reproduce the results of the tight-binding model. This confirms the observation that the larger gap between the two bands at the valley minimum results in less anisotropy of the conduction band.

IV. Shear Strain

In the principal coordinate system uniaxial stress along the [110] direction generates diagonal (ε_{ii}) as well as off-diagonal (ε_{xy}) components of the strain tensor. The diagonal components are added to the diagonal matrix elements (1) of the [001] valley [15]:

$$H_{ii} = H_{ii}^0 + \delta E_C, \quad (6)$$

where $\delta E_C = \Xi_d\,(\varepsilon_{xx}+\varepsilon_{yy}+\varepsilon_{zz}) + \Xi_u\,\varepsilon_{zz}$, with Ξ_d denoting the dilation and Ξ_u the uniaxial deformation potentials for the conduction band. The off-diagonal elements of the Hamiltonian are modified by the shear strain components [9]:

$$H_{ij}(k) = H_{ij}^0 - D\varepsilon_{xy}, \quad (7)$$

where $D = 14$ eV denotes the shear deformation potential.

The dispersion relation of the [001] valleys including the shear strain component for the conduction band now reads:

$$E(\mathbf{k}) = \frac{\hbar^2 k_z^2}{2m_l} + \frac{\hbar^2(k_x^2+k_y^2)}{2m_t} + \delta E_C - \sqrt{\left(\frac{\hbar}{m_0}k_z p\right)^2 + \left(D\varepsilon_{xy} - \frac{\hbar^2 k_x k_y}{M}\right)^2}. \quad (8)$$

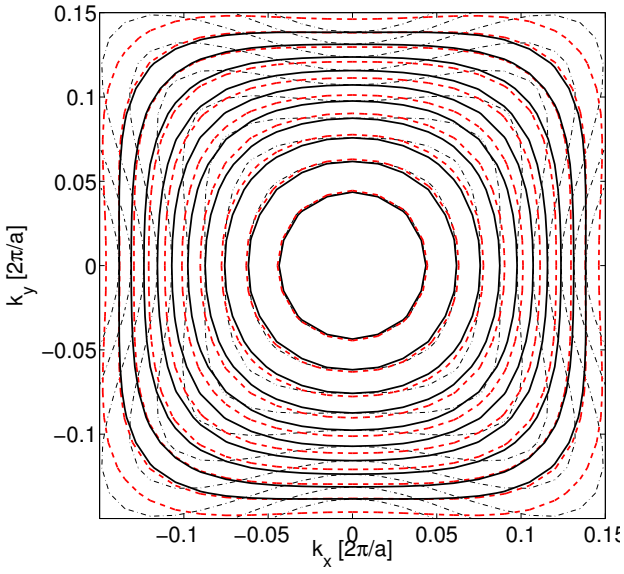

Fig. 5. Dispersion at the valley minimum obtained from the $sp^3d^5s^*$ model (dashed contour lines), from (3) with the correct EPM value Δ=0.53 eV (dotted-dashed contours), and from (3) with Δ=1.2 eV (solid lines). The distance between the equi-energy contour lines is 50 meV.

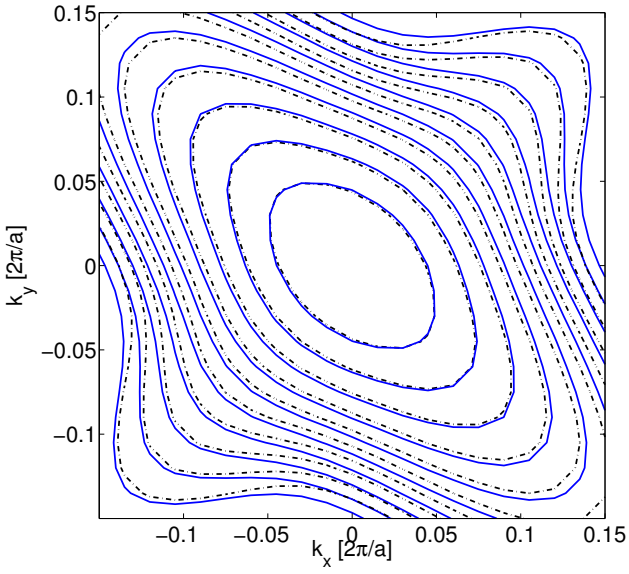

Fig. 6. Comparison between the EPM band structure calculations and (8) at the minimum $k_{\min}$ for shear strain ε_{xy}= 1%. The analytical model predicts the anisotropic dependences of the transversal effective mass on shear strain. The Contours are spaced at 50 meV.

Because of shear strain the minimum $k_{\min}$ moves closer to the X point. From (8) we obtain

$$k_{\min} = -k_0\sqrt{1-\eta^2}, \quad |\eta| < 1 \tag{9}$$

Here, the dimensionless off-diagonal strain $\eta = 2D\varepsilon_{xy}/\Delta$ is introduced. Interestingly, for $\eta \geq 1$ the valley minimum rests exactly at the X point. In Fig. 6 we compare the results of the EPM calculations with the analytical model (8) at the minimum $k_z = -k_{\min}$ for ε_{xy}= 1%. The agreement between the two-band **k**·**p** model and the EPM calculations is excellent. In contrast to the unstrained case shown in Fig. 4 the dispersion relation at small k_x and k_y becomes anisotropic displaying the dependence of the transversal effective mass on shear strain. Evaluating the second derivatives of (8) at the band minimum (9), we obtain two different branches for the effective mass across (m_{t1}) and along (m_{t2}) the stress direction [110]:

$$\frac{m_t}{m_{t1}(\eta)} = \begin{cases} \left(1-\eta\dfrac{m_t}{M}\right) & , \quad |\eta|<1 \\ \left(1-\mathrm{sgn}(\eta)\dfrac{m_t}{M}\right) & , \quad |\eta|>1 \end{cases} \tag{10}$$

$$\frac{m_t}{m_{t2}(\eta)} = \begin{cases} \left(1+\eta\dfrac{m_t}{M}\right) & , \quad |\eta|<1 \\ \left(1+\mathrm{sgn}(\eta)\dfrac{m_t}{M}\right) & , \quad |\eta|>1 \end{cases} \tag{11}$$

Here, $\mathrm{sgn}(\eta)$ denotes the sign function.

V. Conclusion

An efficient two-band **k**·**p** model for the conduction band in silicon is proposed. Predictions of the two-band **k**·**p** model are compared with the results of the empirical nonlocal pseudo-potential method and the $sp^3d^5s^*$ nearest-neighbor tight-binding model. It is demonstrated that the two-band **k**·**p** model is consistent with the results of the empirical pseudo-potential method. The model accurately describes the whole band structure around the valley minimum, including the effective masses and the band non-parabolicity. It is shown that the $sp^3d^5s^*$ model overestimates the gap between the two lowest conduction bands at the valley minimum and consequently underestimates the anisotropy due to non-parabolic effects. The two-band **k**·**p** model allows to account for shear stain, which leads to profound changes in the conduction band causing the valley minima to shift and the effective masses to change. Predictions of the the two-band **k**·**p** method are in good agreement with the results of the pseudo-potential method.

Acknowledgment

This work was supported by the Austrian Science Fund, project P19997-N14.

References

[1] J. M. Luttinger and W. Kohn, Physical Review **97**, 869 (1955).
[2] K. Uchida, A. Kinoshita, and M. Saitoh, in *IEDM Techn. Dig.* (2006), pp. 1019–1021.
[3] K. Rim *et al.*, in *VLSI Symp. Techn. Dig.* (2002), pp. 98–99.
[4] S. I. Takagi, J. L. Hoyt, J. J. Welser, and J. F. Gibbons, J.Appl.Phys. **80**, 1567 (1996).
[5] K. Uchida, T. Krishnamohan, K. C. Saraswat, and Y. Nishi, in *IEDM Techn. Dig.* (2005), pp. 129–132.
[6] E. Ungersboeck, V. Sverdlov, H. Kosina, and S. Selberherr, in *International Conference on Simulation of Semiconductor Processes and Devices* (2006), pp. 43–46.
[7] E. Ungersboeck *et al.*, IEEE Trans.Electron Devices (in print, 2007).
[8] V. Sverdlov, E. Ungersboeck, H. Kosina, and S. Selberherr, in *Proc. EUROSOI 2007* (January, 2007), pp. 39–40.
[9] J. C. Hensel, H. Hasegawa, and M. Nakayama, Physical Review **138**, A225 (1965).
[10] D. Rideau *et al.*, Physical Review B **74**, 195208 (2006).
[11] M. M. Rieger and P. Vogl, Physical Review B **48**, 14276 (1993).
[12] S. Datta, *Quantum Transport: Atom to Transistor* (Cambridge University Press, 2005).
[13] T. B. Boykin, G. Klimeck, and F. Oyafuso, Physical Review B (Condensed Matter and Materials Physics) **69**, 115201 (2004).
[14] C. Jacoboni and L. Reggiani, Reviews of Modern Physics **55**, 645 (1983).
[15] I. Balslev, Physical Review **143**, 636 (1966).