Stress-Induced Valley Splitting in Silicon Thin Films

Viktor Sverdlov, Thomas Windbacher, Hans Kosina, and Siegfried Selberherr

Institute for Microelectronics, TU Wien, Gusshausstrasse 27-29, A-1040 Wien, Austria
{sverdlov|windbacher|kosina|selberherr}@iue.tuwien.ac.at

1. Introduction

Downscaling of MOSFETs is successfully continuing because of innovative changes in the technological process and introduction of new materials. The 45 nm MOSFET process technology recently developed by Intel [1] involves new hafnium-based high-k dielectric and metal gates and represents a major change in the technological process since the invention of the MOSFET. Although alternative channel materials with mobility higher than that in Si were already investigated [2], strained Si is considered as one of the main candidates for channel material beyond the 45 nm technology node.

The band structure and transport properties of Si have been the subject of intensive research for many decades and a satisfactory understanding of many issues has been achieved. However, new technological challenges and engineering solutions constantly reveal shortcomings of previously used models. For example, stress engineering employed to enhance mobility of modern MOSFETs prompted a reconsideration of the conduction band structure of Si in order to include modifications of the effective masses under shear strain [3]. The analytical two-band **k·p** model [4,5] for a conduction band valley describes the observed dependence of the electron effective masses on shear strain. While including the previously used parabolic approximation for the conduction band valleys as a limiting case of energies close to the valley minimum, the two-band **k·p** model accurately describes the dispersion relations for energies up to approximately 0.5 eV.

The two-band **k·p** model provides a more general approach to compute the band structure in thin films. This allows an analysis of the dependence of subband energies and dispersions on film thickness for arbitrary stress conditions. At the same time it reveals shortcomings of the **k·p** method to describe the subband structure in strained Si films. The most obvious of the limitations is its inability to address the splitting between the subbands belonging to equivalent valleys. Indeed, due to the valley degeneracy, the unprimed subbands in (001) Si films remain two-fold degenerate within the **k·p** approach. However, a valley splitting larger than the spin splitting was reported recently [6]. It is necessary to go beyond the **k·p** theory by restoring some important properties of the periodic Bloch amplitude.

The Bloch amplitude is deduced from a tight-binding model with two orbitals per site, which mimics dispersion relations of the underlying two-band **k·p** model. Quantization conditions using the total wave function including the Bloch amplitude describe the energy splitting between the two unprimed subband ladders in an ultra-thin Si film. Lifting degeneracy between the two valleys reduces scattering and improves the coherence time in Si-based spin qubits, which makes silicon-based quantum devices promising for future applications in quantum computing. We demonstrate that the valley splitting in (001) Si films can be controlled and significantly enhanced by applying stress in [110] direction.

2. Subband structure from the two-band k·p model

From symmetry consideration it follows that the two-band **k·p** Hamiltonian of a [001] valley in the vicinity of the X point of the Brillouin zone in Si must be of the form [7]:

$$H = \left(\frac{\hbar^2 k_z^2}{2m_l} + \frac{\hbar^2 (k_x^2 + k_y^z)}{2m_t}\right) I + \left(D\varepsilon_{xy} - \frac{\hbar^2 k_x k_y}{M}\right)\sigma_z + \frac{h^2 k_z k_0}{m_l}\sigma_y, \quad (1)$$

where $\sigma_{y,z}$ are the Pauli matrixes, I is the 2×2 unity matrix, m_t and m_l are the transversal and the longitudinal effective masses, respectively, $k_0 = 0.15 \times 2\pi / a$ is the position of the valley minimum relative to the X point in unstrained Si, ε_{xy} denotes the shear strain component, $M^{-1} \approx m_t^{-1} - m_0^{-1}$, and D=14 eV is the shear strain deformation potential [4,5]. The two-band Hamiltonian results in the following dispersions:

$$E = \frac{\hbar^2 k_z^2}{2m_l} + \frac{\hbar^2 (k_x^2 + k_y^2)}{2m_t} \pm \sqrt{\left(\frac{\hbar^2 k_z k_0}{m_l}\right)^2 + \delta^2}\,, \quad (2)$$

where the negative sign corresponds to the lowest conduction band and

$$\delta^2 = (D\varepsilon_{xy} - \hbar^2 k_x k_y / M)^2 . \quad (3)$$

Note that in (2) all the wave vectors and energies are counted from the X-point of the Brillouin zone. The usual parabolic approximation is obtained from (2), when coupling between the two conduction bands described by the parameter δ is neglected. Coupling between the bands is small, when the wave vectors

978-1-4244-1730-8/08/$25.00 ©2008 IEEE

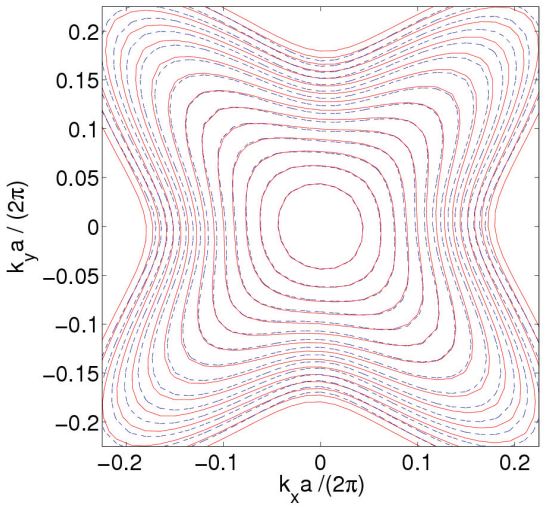

Fig.1: *Comparison between Eq.(2) (dashed lines) and the results of EPM calculations (solid lines). The contour lines are spaced at 50 meV. Tensile stress along [110] and compressive stress along [-110] of 150 MPa in each direction is applied resulting in the only non-zero shear strain component.*

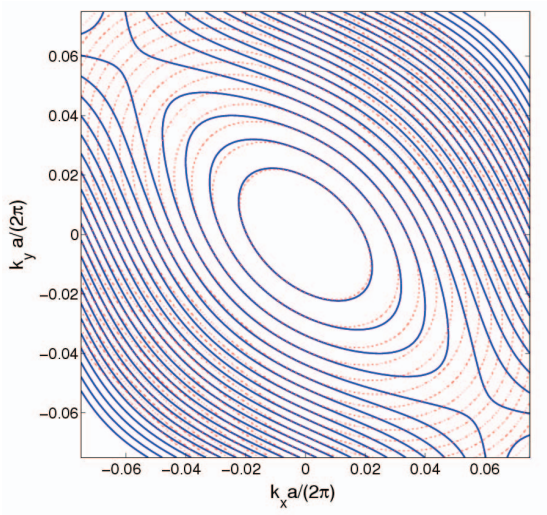

Fig.2: *Comparison between the dispersion Eq.(4) (solid lines) and parabolic approximation with the strain-dependent effective masses (7) for unprimed subbands. t=5.4 nm, $\varepsilon_{xy} = 0.9\%$ The contour lines are spaced at 10 meV. Good agreement is obtained up to 25meV.*

$|k_x|$, $|k_y|$ << $k_0\,(M/m_l)^{1/2}$ and shear strain $\varepsilon_{xy} = 0$. Due to band coupling the dispersion (2) becomes non-parabolic, if the shear strain component is non-zero or the energies are high. In order to check the accuracy of (2) we have carried out numerical band structure calculations using the empirical pseudo-potential method (EPM) with parameters from [8, 9]. Excellent agreement between the two-band **k·p** model (1) and the EPM results is demonstrated in Fig.1. Tensile stress in [110] and compressive stress in orthogonal [-110] direction of 150 MPa is applied resulting in the only non-zero component ε_{xy}. Shear strain ε_{xy} substantially modifies the energy dispersion of the [001] valleys even at small energies making the dispersion relation anisotropic [5, 6]. Shear strain also modifies the k_z dispersion: the valley minimum moves both in energy and position, approaching the *X*-point of the Brillouin zone for larger strain $\varepsilon_{xy} \to k_0^2/(m_l D)$. At the same time the k_z dispersion becomes highly non-parabolic.

The dispersion (2) is valid in a larger range of energies compared to parabolic dispersion with isotropic non-parabolic correction and can be used to determine the subband structure in thin Si films. The subband energies can be found analytically for an infinite potential well, which is a good model for an ultra-thin Si film. The dispersion of the unprimed subbands in a [001] thin Si film of thickness *t* is

$$E_n(k_x,k_y) = E_n^0(k_x,k_y) - \frac{\delta^2 m_l}{2k_0^2(1-p_n^2)}, \qquad (4)$$

where

$$q_n = (\pi n)/(tk_0)$$

and E_n^0 is the subband dispersion for parabolic bands:

$$E_n^0(k_x,k_y) = \frac{\hbar^2\pi^2 n^2}{2m_l t^2} + \frac{\hbar^2(k_x^2+k_y^2)}{2m_t} - \frac{\hbar^2 k_0^2}{2m_l}.$$

Eq.(4) is valid when

$$\left(1-q_n^2\right)^2 > \delta^2 m_l^2/(\hbar^4 k_0^4). \qquad (5)$$

The dispersion (4) describes the subband quantization energy correction due to strain with respect to the valley minimum:

$$\Delta E_n(\varepsilon_{xy}) = -\frac{\pi^2 n^2}{2m_l t^2}\frac{(D\varepsilon_{xy} m_l)^2}{\hbar^2 k_0^4(1-q_n^2)}. \qquad (6)$$

This expression is obtained after taking into account the strained induced valley minimum energy shift $\Delta E(\varepsilon_{xy}) = -(D\varepsilon_{xy})^2 m_l/(2k_0^2)$ and the dependence of the longitudinal mass m_l on strain:

$$m_l(\varepsilon_{xy}) = m_l\left(1-(D\varepsilon_{xy}m_l)^2/\hbar^4 k_0^4\right)^{-1}$$

The dispersion (4) also describes corrections to the transversal mass m_t due to strain, thickness *t,* and subband number *n*:

$$m_n^{\mp} = m_t\left(1 \pm \frac{D\varepsilon_{xy} m_l}{\hbar^2 k_0^2}\frac{m_t}{M}\frac{1}{1-q_n^2}\right)^{-1} \qquad (7)$$

Here m_t^- is the effective mass component along the direction [110] of tensile strain. Comparison of the dispersion relation (4) to the parabolic approximation with transversal masses (7) is demonstrated in Fig.2. Deviations from the parabolic approximation become significant for electron energies above 20 meV, and the dispersion (4) should be used instead of a

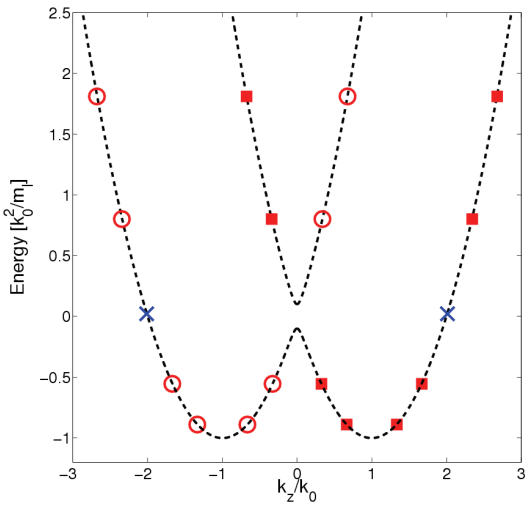

Fig.3: *Subband energies from the two-band* ***k·p*** *theory according to (4 and (6) with p_z determined by (9). Squares indicate the subbands centered around k_0, while circles correspond to those around $-k_0$. Subbands with p_z=0 are shown with crosses. Their energy lies in the gap between the two conduction bands at the X-point.*

parabolic approximation at higher carrier concentrations.

The dispersion (4) is valid only when condition (5) is satisfied. The quantization energies obtained from (4) for non-zero strain are indicated in Fig.3 by circles and allow a simple graphic interpretation. For a fixed film thickness the quantization energy E_n is obtained from the equation:

$$E\left(q_z - \frac{\pi n}{t}\right) = E\left(q_z + \frac{\pi n}{t}\right), \qquad (8)$$

where $E(k_z)$ is determined by (2). In unstrained Si the k_z dispersion is parabolic and $q_z = \pm k_0$. For non-zero strain a gap between the two conduction bands opens at the X-point and the k_z dispersion becomes non-parabolic. It results in the dependence of q_z on strain, thickness and subband number:

$$q_z = \pm k_0 \sqrt{1 - (m_l \delta / \hbar^2 k_0^2)^2 / (1 - q_n^2)}\,. \qquad (9)$$

The two signs in (9) correspond to two sets of subband ladders. We notice that, according to (4,5), within the two-band **k·p** theory the subband energies are the same, so the two subband ladders are *degenerate*. Therefore, the two-band **k·p** theory is unable to describe the splitting between the two subband ladders from different valleys. This splitting was predicted theoretically [10] and confirmed by several numerical calculations [11, 12]. It was recently demonstrated experimentally that the splitting is quite large [6], so one important limitation of the two-band **k·p** theory is that it cannot describe this splitting.

If the condition (5) is not fulfilled, Eq. (8) can be satisfied only with q_z=0. This results in a subband dispersion obtained from (2) with $k_z = \pi n / t$. The solutions shown by crosses in Fig.3 are clearly different from those described above by (9) with non-zero q_z. In any real system, however, the two types of solutions are not completely independent. A smooth transition between them is anticipated, when the parameters are gradually modified. Shear train which regulates the gap and band non-parabolicity is a particular example of a suitable parameter. With the gap increased, a pair of degenerate subbands eventually reaches the maximum of the dispersion relation at the X-point, where it moves into the gap and becomes the subband with p_z=0, marked by crosses in Fig.3. However, the number of subbands at the transition is not conserved and one subband is missing, since there is only a single solution with p_z=0 within the gap. This is another shortcoming of the two-band **k·p** theory. In order to resolve the crossover between the two types of solutions for subband energies accurately and describe the valley splitting correctly one has to go beyond the two-band **k·p** theory.

3. Beyond the two-band k·p model: Controllable valley splitting

We introduce an auxiliary tight-binding model defined on a lattice of sites each containing two localized orbitals α (z) and β (z). The orbitals are defined in such a way that the expressions

$$u(z) = \sum_{n=-\infty}^{\infty} \alpha(z - na/2); \quad v(z) = \sum_{n=-\infty}^{\infty} \beta(z - na/2) \qquad (10)$$

reproduce the z-dependence of the two Bloch functions corresponding to the two lowest conduction bands at the X-point of the Brillouin zone. Then for an arbitrary wave vector k_z the Bloch function is written as

$$\psi(z, k_z) = e^{ik_z z}(a(k_z)u(z) + b(k_z)v(z))\,, \qquad (11)$$

where u corresponds to the lowest band. Using (10), we rewrite the wave function (11):

$$\psi(z, k_z) = \sum_{n=-\infty}^{\infty} e^{ik_z na/2}(a(k_z)\alpha(z - na/2) + ib(k_z)\beta(z - na/2)), \qquad (12)$$

where the argument z in the exponent is approximated by $na/2$ and the real coefficients $a(k_z)$ and $b(k_z)$ are found from the 2×2 Schrödinger equation with the Hamiltonian (1). Strictly speaking, the orbital functions in (11) are related to the ones in (10) as

$$\alpha'(z - na/2) = e^{ik_z(z - na/2)} a(z - na/2)\,.$$

However, we can neglect this difference under the condition $k_z a << 1$ when the two-band Hamiltonian (1) is justified.

There are two pairs of roots $k_z = \pm k_1, \pm k_2$ of the dispersion relation (2) for a particular energy E. Then

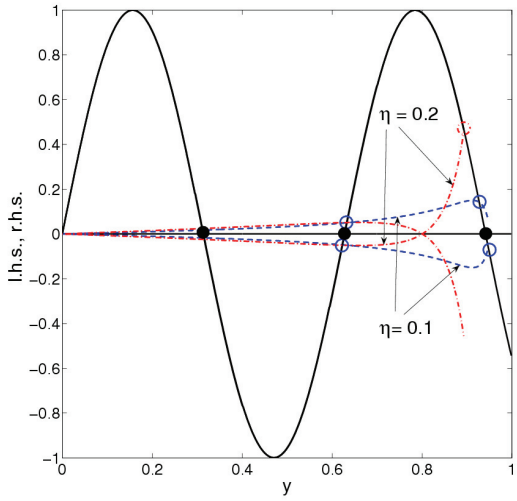

Fig.4: *Graphical solutions of (14). Without strain the roots are two-fold degenerate (filled circles). For non-zero shear strain the degeneracy is lifted, which results in energy splitting (15) between the unprimed subband ladders from different valleys (open circles), which is proportional to strain and can become larger than the spin splitting.*

the wave functions for a finite array of $2N$ sites are:

$$\psi_{\pm}(z,E) = \sum_{i=1,2}\sum_{n=-N}^{N} C_i \exp(ik_i na/2)(a(k_i) \times \alpha(z-na/2) + ib(k_i)\beta(z-na/2)) \pm C.C., \quad (13)$$

where $C.C.$ stands for complex conjugate. In writing (13) we have assumed that both roots $k_1>k_2$ are real. If the root k_2 is imaginary, the coefficient $b(k_2)$ is also imaginary. This corresponds to the solutions shown in Fig.3 by crosses. Therefore, dispersion relations based on (13) are able to describe a smooth transition between the two types of the solutions in Fig.3, when strain is increased.

Dispersion relations are obtained by setting the wave function (13) to zero at the film interfaces $z=\pm t/2$ [11]. Introducing a new variable $y_n = (k_1 - k_2,)/k_0$, we obtain for it the following equation:

$$\sin(y_n k_0 t) = \pm \frac{\eta p_n^{\pm} \sin\left(\frac{1-\eta^2-y_n^2}{1-y_n^2} k_0 t\right)}{\sqrt{(1-y_n^2)(1-\eta^2-y_n^2)}}, \quad (14)$$

where dimensionless strain $\eta = m_l D\varepsilon_{xy}/\hbar^2 k_0^2$ is introduced. This equation, the main result of the paper, is briefly analyzed below.

It follows from (14) that for $\delta = 0$, when the dispersion is purely parabolic, the right-hand side (r.h.s.) of (14) is zero. Therefore, for a parabolic dispersion the valley splitting is exactly zero for any film thickness. For $\eta \neq 0$ the r.h.s is shown in Fig. 4, together with the left-hand side of (14). One sees that due to the plus/minus sign in the r.h.s a two-fold degenerate solution p_n splits into two solutions for a non-parabolic band. For small η the solution is in the form $y_n = p_n \pm \zeta$. It leads to the valley splitting

$$\Delta E_n = 2\left(\frac{\pi n}{k_0 t}\right)^2 \frac{D\varepsilon_{xy}}{k_0 t}\sin(k_o t) \cdot \quad (15)$$

In accordance with earlier publications [10, 11], the valley splitting is inversely proportional to the third power of k_0 and the third power of the film thickness t. The value of valley splitting oscillates with film thickness, in accordance with [11]. In contrast to previous work the subband splitting is proportional to the gap δ at the X-point, and not at the Γ-point. The parameter δ, which determines non-parabolicity, depends strongly on shear strain. Application of uniaxial [110] stress to [001] ultra-thin Si film generates valley splitting proportional to strain. Therefore, the valley splitting (15) can be controlled by adjusting strain and thickness t. Currently, uniaxial stress is already used for performance enhancement of modern MOSFETs, where it is introduced in a controllable way, and its application to control valley splitting is technologically possible.

4. Conclusion

In conclusion, an alternative way to induce controllable valley splitting in ultra-thin Si films by introducing shear strain generated by [110] uniaxial stress is proposed. For small stress values the splitting is shown to depend linearly on shear strain. Valley splitting rapidly increases with decreasing Si thickness and can be larger than the spin splitting.

5. Acknowledgment

Financial support from the Austrian Science Fund FWF trough the Project P19997-N14 is gratefully acknowledged.

References

[1] K.Mistry *et al.*, IEDM 2007, p.247.

[2] M.K.Hudati *et al.*, IEDM 2007, p.625.

[3] K.Uchida *et al.*, IEDM 2005, p.129

[4] J.C. Hensel *et al.*, *Phys.Rev.,* **138**, p.A225, 1965.

[5] V.Sverdlov *et al.*, ESSDERC 2007, p.386.

[6] S. Goswami *et al.*, Nature Physics, **31**, p.41, 2007.

[7] G.L.Bir and G.E.Pikus, Symmetry and Strain-Induced Effects in Semiconductors, Nauka, Moscow 1972 (English translation J.Willey&Sons, NY 1974).

[8] M. Rieger and P. Vogl, *Phys.Rev.B,* **48**, p.14275, 1993.

[9] E. Ungersboeck *et al.*, *IEEE T-ED*, **54**, 2183, 2007.

[10] T.Ando, A.B.Fowler, F.Stern, *Rev.Mod.Phys.*, **54**, 437, 1982.

[11] T.B. Boykin *et al.*, *Phys.Rev.B,* **70**, 165325, 2004.

[12] D.Esseni and P.Palestri, *Phys.Rev.B,* **72**, 165342, 2005.